PhD thesis

Louise Christiane Rosenberg Christ

Downstream events of neurogenin 3 in pancreatic endocrine differentiation

Academic advisor: Professor Olaf Nielsen

Submitted: 15/10/09
Preface

This thesis is based on experiments performed in the Department of Developmental Biology at the Hagedorn Research Institute, Gentofte, from October 2004 to September 2009 including two periods of maternity leave of ten months each. Supervisors on the project were Professor Olaf Nielsen, Functional Genomics, Institute of Biology, Faculty of Science, University of Copenhagen, and Jacob Hecksher-Sørensen, Ph.D., staff scientist, Department of Developmental Biology, Hagedorn Research Institute. The thesis is submitted to meet the requirements for obtaining a Ph.D. degree at the Institute of Biology, Faculty of Science, University of Copenhagen.

One manuscript and two published papers are included:

Louise C Rosenberg, Merete L Lafon, Hani Yassin, Jan Nygaard Jensen, Jacob Hecksher-Sørensen, Palle Serup. The transcriptional activation potential of Neurog3 affects migration and differentiation of ectopic endocrine cells in the chicken endoderm. (Manuscript in preparation, appendix 1).


Cover pictures:
Left: Early chicken gut endoderm electroporated with EGFP, stained for Foxa2 (grey), GFP (green), and glucagon (red).
Middle: Early chicken gut endoderm electroporated with Neurog3^{74-214}, stained for Foxa2 (grey), GFP (green), and insulin (red).
Right: 3T3 cells transfected with Neurog3^{91-214}, stained for HA-tag (red), GFP (green), and DAPI (grey) (upper panel), or with Myt1 6- and 7-zinc finger, stained for HA-tag (red), myc-tag (blue), and GFP (green) (lower panel).
Acknowledgements

I would like to thank head of department Helle V. Petersen, Palle Serup and my supervisor Jacob Hecksher-Sørensen for giving me the opportunity to do a project on developmental biology at Hagedorn Research Institute, and the rest of Department 474 for giving me such a good time. It has been a real pleasure to be a part of Hagedorn all this time, and to get to know all of you. Especially, thanks to Jacob Hecksher-Sørensen for inspiration, input, help, and discussion. Also thanks to my University supervisor Olaf Nielsen.

A great thank goes to Hanne Duus Laustsen for her patience and help on any matter in the laboratory! Also thanks to Tove Funder Larsen for the work she did with $^{32}$P when I was expecting Solvej. A special thank goes to Hani Yassin for his friendship, for being such a funny guy, and for all the talks we had in the lab. I believe Jacob still owes him a dinner for providing the very special combination of “anti-p300” and the 14-amino acid extended version of Neurog3 that gave so interesting results!

Finally, I would like to thank my friends and family for listening so patiently to the endless amount of details on EMSAs and eggs. Especially, thanks to Jan Van Deurs for actually wanting all my pictures of cells for a screen saver!

And lastly, a huge thank to Egil for being supportive and understanding, and to Sander and Solvej for putting everything into perspective.

Louise C Rosenberg Christ,
Gentofte, September 2009
Summary

Diabetes affects more than 170 million people worldwide and a prognosis suggests that the prevalence will increase over the next 30 years. In Denmark, the prevalence doubled from 1996 to 2007, where 240,000 people were affected. Diabetes is the manifestation of abnormalities regarding glucose homeostasis which may have several causes. The most prevalent types are a result of either an auto-immunological destruction of the insulin producing β-cells (type 1) or a combination of insulin resistance of the tissues and impaired β-cell function (type 2) in both cases leading to hyperglycaemia.

Research in the developmental biology of the pancreas and the mechanism behind differentiation of the pancreatic progenitor cells into mature β-cells provide knowledge of the molecular events and signals of the process that may be applied to stem cell research. The aim is to be able to induce directed differentiation of stem cells into insulin producing cells in vitro. The hope is that successful production of β-cells from stem cells combined with a good protocol for the transplantation of these cells into human beings would offer an opportunity for diabetes patients to obtain a normal glucose homeostasis and be relieved of the eventually debilitating long-term effects of diabetes and the trouble with insulin injections several times a day.

In pancreatic development the expression of the transcription factor Neurog3 is a key event, as it marks the transition from the stage in development where the pancreas progenitor cells proliferate to form the cell pool of the future organ to the stage in development where the endocrine precursor cells and the individual endocrine cell types are specified. Neurog3 specifically induces endocrine differentiation that eventually leads to the generation of the mature hormone producing cells of the islets of Langerhans, and among these are the insulin producing β-cells. The function of Neurog3 as a transcription factor and an endocrine inducer is well established, but the immediate events regarding Neurog3 function are poorly understood.

The scope of this thesis was to investigate these immediate functions of Neurog3 in endocrine development. We applied a deletion analysis to Neurog3 and investigated the resulting truncated Neurog3 proteins with regard to their cellular localisation, DNA-binding capacity, activity and in vivo function applying the technique of in ovo electroporation. Furthermore, the role of the binding partner E12 was investigated with regard to Neurog3 function. Mutations in the human Neurog3 were investigated with regard to their DNA binding capacity and the effect of E12 on their function in vivo. Finally, the transcription factor Myt1 was investigated as a putative direct target of Neurog3.
The truncated Neurog3 proteins revealed that an intact bHLH domain and the region of amino acids flanking the bHLH on the N-terminal side may be involved in nuclear localisation as the proteins lacking the N-terminal domain or with a truncated bHLH domain were present to a larger extent in the cytoplasm.

The N-terminal domain was not required for Neurog3 to induce activation of the Neurod1 promoter to the same level as the wild type protein, and the protein lacking the N-terminal domain had a very good DNA binding capacity. Truncations in the C-terminal domain reduced the transcriptional activity of Neurog3 on the Neurod1 promoter. In vivo the truncated proteins showed that even low levels of Neurog3 activity induced differentiation of glucagon expressing cells, migration and clustering. Surprisingly, the lack of the N-terminal domain enabled Neurog3 to induce ectopic insulin expression in vivo in the endoderm of the early chicken gut tube endoderm. This suggests that Neurog3 may be involved in lineage specification in addition to the function as an inducer of general endocrine fate. However, further investigations are needed to support this hypothesis.

The Neurog3 binding partner E12 did not enhance Neurog3 transcriptional activity on the Neurod1 promoter and did not increase Neurog3 induced differentiation in vivo in early chicken gut tube endoderm. This was surprising, and it is unknown if this observation is a direct effect of E12 or if it is a result of Neurog3 independent effects of E12 in the endodermal cells. However, the observation that Neurog3 induced migration but not differentiation when co-electroporated with E12 supports previous findings that the processes of Neurog3 induced differentiation and migration are uncoupled. The same effect of E12 was observed when it was co-electroporated with one of the human Neurog3 mutants or the double mutant. These mutations were found to have reduced DNA binding.

Finally, Myt1 was shown to be a direct target of Neurog3, as it induced transcription from the Myt1 promoter. The truncated Neurog3 proteins showed a slightly different effect on the Myt1 promoter as the N-terminal domain was required for Neurog3 to induce a full transcriptional response, and truncations from the C-terminal end also seemed to have a greater impact on the response of the Myt1 promoter than of the Neurod1 promoter. The Myt1 protein variants were located to different cellular compartments although they were able to interact. The impact of this observation is unknown.

In conclusion, this thesis supports that Neurog3 induced endocrine differentiation is a diverse process that involves transcription of several target genes perhaps in a differential way, and
that Neurog3 mediated differentiation and migration are uncoupled processes. It is also shown that the transcription factor Myt1 is a direct target of Neurog3. An intact bHLH domain and the amino acids flanking the bHLH domain on the N-terminal side are involved in nuclear localisation. It was unexpected that E12 did not enhance Neurog3 mediated differentiation or \textit{in vitro} activity. The ability of the Neurog3 protein lacking the N-terminal domain to bind DNA more efficiently and induce ectopic insulin expression suggests that further investigations of Neurog3 may reveal nuances of regulation and function that may be of relevance for the application of developmental biology to directed differentiation of stem cells into insulin expressing cells.
Resumé

Mere end 170 millioner mennesker har diabetes og en prognose viser at dette tal forventes at stige over de næste 30 år. I Danmark er forekomsten fordoblet fra 1996 til 2007, hvor 240.000 mennesker havde diabetes. Diabetes er manifestationen af en unormal blodsukker regulering, som kan have flere årsager. De hyppigst forekommende typer skyldes enten en autoimmun ødelæggelse af de insulin-producerende β-celler i bugspytkirtlen (type 1) eller en kombination af insulin resistens i vævene og en utilstrækkelig β-celle funktion (type 2). I begge tilfælde kommer det til udtryk som forhøjet blodsukker.

Viden om hvordan bugspytkirtlen dannes og om mekanismerne bag differentieringen af bugspytkirtlens tidlige celler til modne hormonproducerende β-celler håber man at kunne overføre til stamcelleforskningen. Det er målet at blive i stand til at inducere stamceller til at differentiere til insulin producerende celler på en kontrolleret måde. Håber er, at en produktion af β-celler i kombination med en protokol for transplantation af disse til mennesker vil give diabetes patienter mulighed for at opnå normal blodsukkerregulering, mindske risikoen for senkomplikationer og at frigøre dem for insulin indsprøjtningerne flere gange om dagen.


Undersøgelserne af de trunkerede Neurog3 proteiner viste, at et intakt bHLH domæne og de aminosyrer, der flankerer bHLH domænet på den N-termiale side, har betydning for kernelokalisering af Neurog3, da de proteiner der manglende den N-terminal ende eller med et trunkeret bHLH domæne var til stede i cytoplasma i højere grad end vildtypen.
Det N-terminal domæne var ikke nødvendigt for at Neurog3 kunne inducer transskription fra Neurod1 promotoren i samme grad som vildtypen, og proteinet der manglede det N-terminale domæne bandt bedre end vildtypen til DNA. Trunkering fra den C-terminale ende reducerede Neurog3s transskriptionsaktivitet fra Neurod1 promotoren. Selv Neurog3 proteinerne med meget lav in vitro aktivitet var i stand til at inducere differentiering af glukagonproducerende celler, migration og sammenklumpning in vivo. Det Neurog3 protein, der manglede det N-terminale domæne, var overraskende nok i stand til at inducere ectopisk insulinproduktion i endodermens af kyllingefostre. Dette kunne tyde på, at Neurog3 muligvis er involveret i specifikationen af de forskellige endokrine celltyper ud over at inducere en generel endokrin skæbne i cellerne. Flere undersøgelser kræves dog for at underbygge denne hypotese.


Myt1 blev identificeret som et direkte mål for Neurog3 transskription. Transkriptionsprofilen af de trunkerede Neurog3 proteiner var lidt anderledes fra Myt1 promotoren end fra Neurod1 promotoren, idet den N-terminale ende var nødvendig for at opnå samme transkriptionsniveau som for vildtypen. Det tydede også på, at trunkeringer fra den C-terminal ende havde en større effekt på responsen fra Myt1 promoteren end fra Neurod1 promoteren. Selvom de var i stand til at interagere, var de to Myt1 protein varianter fortrinsvis lokalisert forskellige steder i cellen, med 6-zink finger formen i cytoplasm og 7-zink finger formen i kernen.

Undersøgelserne beskrevet i denne afhandling tyder på at Neurog3 har flere forskellige funktioner, og støtter de tidligere beskrevne observationer at Neurog3 inducerer migration og differentiering uafhængige processer. Neurog3 inducerer transskription af flere forskellige gener, også af transkriptionsfaktoren Myt1, som er et direkte mål for Neurog3. Det var uventet at bindingspartneren E12 ikke øgede Neurog3 medieret differentiering in vivo eller transkriptionel aktivitet in vitro. Et intakt bHLH domæne og aminosyrer i den region, der flankerer bHLH domænet på den N-terminale side er involverede i kernelokalisering. Det
trunkerede Neurog3 protein, der mangler det N-terminale domæne, binder mere effektivt til DNA end Neurog3 vildtypen og kan inducere ectopisk insulinproduktion i endodermen af den tidlige tarm i kyllingefostre. Disse observationer kunne indikere at Neurog3 kan påvirke differentiering af de forskellige endokrine cellelinier udover at inducere en generel endokrin skæbne. En mere detaljeret forståelse af reguleringen og funktionen af Neurog3 vil give os en viden, der muligvis kan udnyttes i forbindelse med, at man ønsker en kontrolleret differentiering af stamceller til insulin producerende celler, der kan bruges til transplantering.
Abbreviations

Arx: aristateless related homeobox
bHLH: basic-helix-loop-helix
Dll1: delta-like 1
Dorsal: Towards the back
Hes1: hairy and enhancer of split 1
Fgf10: fibroblast growth factor 10
Foxa2: forkhead box A2 (also known as HNF3β)
Gdf11: growth differentiation factor 11
HH: Hamburger and Hamilton, staging of chicken embryos (Hamburger V, 1951)
HLH: helix-loop-helix
Hnf1a: HNF1 homeobox A
Hnf1b: HNF1 homeobox B
Insm1: insulinoma-associated 1 (also known as IA1)
IRES: internal ribosom entry site
Ltbp1: latent transforming growth factor beta binding protein 1 (also known as TGFβ)
Mafa: v-maf musculoaponeurotic fibrosarcoma oncogene family, protein A
Mafb: v-maf musculoaponeurotic fibrosarcoma oncogene family, protein B
Myt1: myelin transcription factor 1
Neurod1: neurogenic differentiation 1
Neurog2: neurogenin 2
Neurog3: neurogenin 3
Nkx2-2: NK2 transcription factor related, locus 2
Nkx6-1: NK6 homeobox 1
Notch1: Notch gene homolog 1
Onecut1: one cut domain, family member 1 (also known as HNF6)
Pax4: paired box gene 4
Pax6: paired box gene 6
Pdx1: pancreatic and duodenal homeobox 1
Ptf1a: pancreas specific transcription factor, 1a (also known as PTF1-p48)
PP: Pancreatic polypeptide
RBPJ: recombination signal binding protein for immunoglobulin kappa J region
Sox9: SRY-box containing gene 9
Ventral: Towards the front
Table of contents

1. Introduction ........................................................................................................... 1
   1.1 The pancreas ........................................................................................................ 1
   1.2 Diabetes ............................................................................................................ 2
   1.3 Stem cell therapy ................................................................................................ 4
   1.4 Developmental biology of the pancreas .............................................................. 5

1.5 Neurog3 is a key transcription factor in endocrine development ....................... 7
   1.5.1 Basic-helix-loop-helix transcription factors .................................................... 7
   1.5.2 Expression of Neurog3 .................................................................................. 10
   1.5.3 Neurog3 function .......................................................................................... 11
   1.5.4 Regulation of Neurog3 .................................................................................. 14
   1.5.4 Neurog3 downstream targets ......................................................................... 17
   1.5.5 The function of Neurog3 target genes and other selected transcription factors ...... 20
   1.5.6 The Neurog3 downstream target Myt1 ......................................................... 23

2. Objectives ............................................................................................................ 25
   2.1 Neurog3 protein domains .................................................................................. 26
   2.2 The role of E12 in Neurog3 function .................................................................. 27
   2.3 Neurog3 mutants in human ................................................................................. 27
   2.4 Neurog3 as the activator of Myt1 transcription ................................................... 28

3. Neurog3 protein domains ..................................................................................... 29
   3.1 Mutation strategy .............................................................................................. 29
   3.2 Expression of the truncated Neurog3 proteins in cells ....................................... 30
   3.3 Cellular localization of the truncated Neurog3 proteins .................................... 32
   3.4 Cellular localisation of the truncated Neurog3 proteins in the presence of wild type Neurog3 ............................................................................................................. 33
   3.5 DNA binding of the truncated Neurog3 proteins ................................................. 35
   3.6 In vitro activity of the truncated Neurog3 proteins ............................................. 39
   3.7 In vivo activity of the truncated Neurog3 proteins ............................................. 41
   3.8 Neurog3 mediated hormone expression .............................................................. 45

4. The role of E12 in Neurog3 function ...................................................................... 47
   4.1 The role of E12 in the cellular localisation of Neurog3 ....................................... 47
   4.2 In vivo activity of Neurog3 in the presence of exogenous E12 .............................. 49
   4.3 In vitro activity of Neurog3 in the presence of exogenous E12 ............................. 52

5. Studies of the Neurog3 mutants discovered in humans ......................................... 55
   5.1 DNA binding of the human Neurog3 mutants .................................................... 55
   5.2 In vivo activity of the human Neurog3 mutants in the presence of E12 .............. 57

6. Neurog3 as an activator of Myt1 transcription ....................................................... 59
   6.1 Investigations on the Myt1 protein isoforms “6-zinc finger” and “7-zinc finger” .... 59

XV
6.2 In vitro response on the Myt1 promoter of the truncated Neurog3 proteins.............. 61
6.3 The effect of exogenous E12 on Neurog3 mediated transcription of Myt1 ............... 63

7. Methods.............................................................................................................................................. 65
7.1 Generation of the Neurog3 mutants................................................................................................. 65
7.2 Gateway cloning .............................................................................................................................. 67
7.3 In vitro generation of proteins ....................................................................................................... 68
7.4 EMSA .................................................................................................................................................. 69
7.5 Cell cultivation.................................................................................................................................... 72
7.6 SDS-PAGE and Western blotting.................................................................................................... 73
7.7 Immunoprecipitation ....................................................................................................................... 74
7.8 Immunofluorescence staining of cells ............................................................................................ 75
7.9 Luciferase assays ............................................................................................................................ 75
7.10 In ovo electroporation .................................................................................................................. 78
7.11 Immunofluorescence staining of whole mounted chicken embryos........................................ 78
7.12 Confocal microscopy of immunofluorescence stained chicken embryos ................................. 79
7.13 Quantification of differentiation .................................................................................................. 79
7.14 Statistics........................................................................................................................................... 80

8. Discussion ........................................................................................................................................... 81
8.1 Neurog3 protein domains ............................................................................................................... 81
8.2 The role of E12 in Neurog3 function ............................................................................................... 85
8.3 Studies of the Neurog3 mutants discovered in humans ............................................................... 87
8.4 Neurog3 as an activator of Myt1 transcription ............................................................................. 87
8.5 Conclusions ....................................................................................................................................... 89

9. References .......................................................................................................................................... 90

Appendix 1. Manuscript 1
Louise C Rosenberg, Merete L Lafon, Hani Yassin, Jan Nygaard Jensen, Jacob Hecksher-Sørensen, Palle Serup. The transcriptional activation potential of Neurog3 affects migration and differentiation of ectopic endocrine cells in the chicken endoderm. (Manuscript in preparation).

Appendix 2. Co-authorship publication 1

Appendix 3. Co-authorship publication 2
1. Introduction

Much research has been carried out to improve the understanding of how the insulin producing cells of the pancreas are specified. The hope is to be able to apply the knowledge from this process to stem cells and produce new β-cells for transplantation to diabetic patients. The hypothesis is that transplantation of β-cells would improve blood glucose control and quality of life and decrease the risk of chronic complications as retinopathy, nephropathy, and neuropathy that eventually affect many diabetic patients. If producing β-cells from stem cells and transplantation would succeed this would be a major achievement in diabetes treatment, and at best a cure.

1.1 The pancreas

The pancreas is located in the abdomen beneath the stomach. It is an organ of exocrine and endocrine tissue. The exocrine cells make up the bulk of the pancreas and secrete a number of digestive enzymes. The epithelial cells lining the ducts in the bottom of the duct system secretes bicarbonate ions. The ducts converge into the pancreatic duct that empties out in the duodenum just after joining the common bile duct from the liver. The digestive enzymes and bicarbonate ions are secreted after a meal as a response to cholecystokinin and secretin released to the blood stream from endocrine cells of the small intestine. The signal initiating this is the presence of amino acids and fatty acids or acid in the small intestine, respectively (Vander AJ, 1994). The endocrine tissue is arranged as islets embedded in the exocrine tissue. These are the islets of Langerhans and they consist of several endocrine cell types; α-cells producing glucagon, β-cells producing insulin, δ-cells producing somatostatin, PP-cells producing pancreatic polypeptide, and ε-cells producing ghrelin. In brief, somatostatin inhibits the release of growth hormone, glucagon and insulin; pancreatic polypeptide suppresses food intake and gastric emptying; ghrelin stimulates growth hormone release and is involved in regulations of appetite; and insulin and glucagon are involved in glucose homeostasis (Rang HP, 1999; Kojima et al., 2007; De Vriese and Delporte, 2008). Insulin is released from the β-cells upon elevated blood glucose levels. Insulin mediates uptake of glucose in muscle, liver, and adipose tissue where it is used for energy supply through glycolysis or energy storage through glycogen or triacylglycerol synthesis. Insulin removes glucose from the blood stream and stabilises the blood glucose levels at times when these are high. Glucagon is released from α-cells and stimulates glycogen breakdown and gluconeogenesis in the liver. This mobilises glucose and stabilises blood glucose values. Glucagon also mediates lipolysis in liver and fat.
cells and catabolism of protein in muscle. The actions of insulin and glucagon function to secure the supply of energy for cells all over the body (Rang HP, 1999).

1.2 Diabetes

Several diseases are related to the pancreas, and diabetes mellitus is one of them. Diabetes mellitus is a metabolic syndrome that is characterised by hyperglycaemia resulting from defects in insulin secretion, insulin action or both. Diabetes is divided into several groups of disease. Type 1 diabetes is characterised by the insufficiency of the β-cells to produce the amount of insulin needed. Type 2 diabetes is characterised by the insufficiency of the tissues to respond to insulin often combined with an insufficient insulin secretion from the β-cells. Other types of diabetes include maturity-onset diabetes of the young (MODY), which is caused by a genetic defect, and gestational diabetes, that affect a fraction of pregnant women (WHO). Type 1 and 2 are the major types.

Type 1 diabetes usually manifests in childhood or adolescence. It is a chronic, autoimmune disease with the destruction of β-cells leading to insufficient insulin production. The acute symptoms are diuresis and hyperglycaemia, and in severe cases ketoacidosis or even coma and death. The disease is treated with insulin injections to account for the lack of endogenous insulin. Blood glucose monitoring and education in the signs and symptoms of hypo- and hyperglycaemia is recommended as tools for better regulation of blood glucose levels. Long-term complications such as atherosclerosis, retinopathy, neuropathy, and nephropathy and often leads to impaired eye-sight, impairment of sensation and increased risk of ulcers and amputations, impaired kidney function, and increased risk of cardio-vascular disease (Rang HP, 1999). A well regulated blood glucose level is believed to delay the onset of these long-term complications (ADA, 1993; Santiago, 1993).

Type 2 diabetes is a syndrome that is usually associated with obesity, unhealthy diet, and decreased physical activity. Hyperglycaemia results from impaired insulin sensitivity or insulin resistance in combination with inadequate insulin release. The patients often acquire impaired glucose tolerance or impaired fasting glycaemia as the syndrome progress. Treatment is focused on changes in life style towards a more healthy diet and increased physical activity, and weight reduction for obese patients. In some instances this is enough to stop or delay further disease. If life style changes are not enough, oral hypoglycaemic agents can be administered. Metformin is such an agent that increases the uptake of glucose in skeletal muscle and inhibits glucose absorption from the intestines and glucose release from the liver.
Sylphonylureas lower blood glucose by stimulating insulin release, but requires a functional β-cell mass (Rang HP, 1999). Increasing insulin release from β-cells is also one of the functions of GLP-1 analogues (Chang et al., 2003). Eventually, many type 2 diabetes patients will depend on insulin administration and many acquire the mentioned long-term complications.

Diabetes is an increasing problem in most age groups of adults all over the world (figure 1.1 B). A conservative 30-year world wide prediction shows that the prevalence is expected to double from the estimated 171 million patients in 2000 up to 366 million patients in 2030 (Wild et al., 2004). The human and economic cost of this will be enormous and it is prioritised by the World Health Organisation (WHO) by the Diabetes Unit and Diabetes Programme (WHO1, WHO2) and by the UN by a resolution from 2006 (UnitedNations, 2007). In Denmark the prevalence doubled from 113,577 patients in 1996 to 240,358 in 2007 (Sundhedsstyrelsen, 2009) (figure 1.1 A).

Figure 1.1. The prevalence of diabetes. A. Development in diabetes prevalence in Denmark 31st December 1996 to 31st December 2007 (modified from (Sundhedsstyrelsen, 2009)). B. Estimates of world wide diabetes prevalence in 2000 and 2030 (modified from (Wild et al., 2004)).
1.3 Stem cell therapy

Even though diabetes treatment and insulin analogues are better these days than ever it is hard to imagine that external means to regulate blood glucose homeostasis will ever be as good as the natural regulation in healthy subjects. Research and initiatives in prevention and treatment are needed to improve the individual patient’s health and to encounter an increasing economical burden. Replacement therapy with pancreas or islet transplantation would pose a potential cure to diabetes. Transplantation of whole pancreas showed generally good results with regard to glycaemic control, improved lipid profiles etc., and a graft survival of 60-80% after 3 years (reviewed in (White et al., 2009). Islet transplantation through the portal vein was performed to type 1 diabetes patients using the so-called “Edmonton-protocol” and endogenous insulin production and glycaemic stability was restored, however, only for a while as the effect was not sustainable (Shapiro et al., 2000; Shapiro et al., 2006). The strategies to transplant whole pancreases or islets both face the problem of donor shortage.

At the emergence of the stem cell therapy field the idea was brought up that the insufficient amount or function of β-cells in diabetic patients could be supplemented with newly produced β-cells generated from stem cells. The hope was that patients would regain their ability to maintain a normal blood glucose homeostasis and be relieved from the insulin injections, and that the risk of long-term complications would be reduced. Stem cell therapy faces several challenges that should be considered. As a continuous source of stem cells would be needed for new patients and maybe several treatments of the same patient, the cells need to be expanded extensively which requires that they are kept in an undifferentiated stage perhaps for many passages. The stem cells will probably have to be differentiated to the final cell type in vitro in order to minimise the risk of uncontrolled growth and cancer. The specificity of the protocol inducing the specific differentiation pathway must be reliable in order to obtain a pure cell population in the final preparation. Finally, for the grafted cells to survive in the new host perhaps some kind of encapsulation of the cells or immunosuppression of the patient would have to be supplied.

A lot of research has been put into all of this and successful attempts to differentiate human embryonic stem cells into definitive endoderm have been reported (D'Amour et al., 2005). It was also reported possible to differentiate the definitive endoderm cells to insulin-expressing cells through pdx-1 and neurogenin 3 (Neurog3) expressing intermediates, and the resulting cells secreted insulin in response to glucose when grafted into mice (D'Amour et al., 2006; Kroon et al., 2008). The fact that the cells going through this differentiation protocol expressed markers that were also identified in human embryonic pancreas development supported the
idea that these cells were endocrine insulin-producing cells that would act as such in a human host, although this remains to be investigated. Recently, the successful induction of human somatic cells to pluripotency and \textit{in vivo} reprogramming of adult pancreatic exocrine cells to \(\beta\)-cells, suggest a strategy that would overcome many of the challenges related to stem cells therapy and the ethical problems of using embryonic cells (Takahashi et al., 2007; Yu et al., 2007; Park et al., 2008; Zhou et al., 2008; Collombat et al., 2009). However, the directed differentiation of induced pluripotent stem cells would still require the application of knowledge obtained from studying embryonic pancreas differentiation.

1.4 Developmental biology of the pancreas

The pancreas develops from the endoderm of the primitive gut tube. A dorsal and a ventral bud appear at e9 in mice and the transcription factor pdx-1 is expressed in the bud cells. The buds grow as the pancreatic progenitor cells proliferate and at e13.5 they have fused to form one organ. The epithelium continue to grow and branch, and acini and ducts become visible as structures at about e14.5 in mouse, whereas endocrine cells continue to be largely individual and associated with the ducts until the end of gestation around e18.5 where they have aggregated into the islets of Langerhans (reviewed in (Slack, 1995; Jorgensen et al., 2007) (figure 1.2). The early period in pancreas formation is often referred to as the primary transition where the cells forming the organ are first determined (Rutter et al., 1968; Jorgensen et al., 2007). The late period in pancreas formation is dominated by specification and differentiation of the endocrine cell types. This stage is often referred to as the secondary transition and in mice it takes place around e13.5-e16.5 (Rutter et al., 1968; Jorgensen et al., 2007). During this period of time some central events take place; the initial definition of the pancreas region in the early gut tube endoderm, the growth and development of the pancreatic progenitor cell pool, and the differentiation of the progenitor cells into mature hormone expressing cells.

The transcription factor Pdx-1 defines the early pancreas region in the gut, and is necessary for pancreas development. Mice lacking pdx-1 selectively lacked the pancreas and ectopic Pdx-1 expression induced migration of the cells and down regulation of genes expressed in the surrounding mesenchyme, but was not sufficient to generate hormone producing cells (Jonsson et al., 1994; Grapin-Botton et al., 2001).

Notch gene homolog 1 (Notch1) signalling is important for progenitor cell proliferation. Small and poorly branched pancreatic buds, accelerated endocrine differentiation, and increased
expression of the transcription factor Neurog3 was observed in mice lacking the Notch1 signalling members delta-like 1 or recombination signal binding protein for immunoglobulin kappa J region (RBPJ) or the Notch1 downstream target Hes-1, a phenotype similar to that observed in mice with early expression of Neurog3 under the pancreatic and duodenal homeobox 1 (Pdx1)-promoter (Apelqvist et al., 1999; Jensen et al., 2000b).

Neurog3 expression terminates progenitor cell proliferation and induces endocrine differentiation. Neurog3 is required for development of pancreatic endocrine cells, as mice lacking Neurog3 fail to develop any pancreatic endocrine cells, and lineage tracing showed that Neurog3 had been expressed in all endocrine cell types (Gradwohl et al., 2000; Gu et al., 2002; Schonhoff et al., 2004). Ectopic expression of Neurog3 was sufficient to induce migration, clustering and differentiation of glucagon- and somatostatin expressing cells in early chicken gut tube endoderm (Grapin-Botton et al., 2001). These functions of Neurog3 makes it an extremely important transcription factor in pancreas development, and understanding Neurog3 function and regulation will help us understand how the progenitor cells differentiate into endocrine hormone producing cells.

**Figure 1.2. Pancreas development.**

A. The dorsal and ventral pancreas buds emerge from the primitive gut tube endoderm, they grow as cells proliferate and later fuse into one organ. Green marks Pdx1 expression. dp: dorsal pancreas bud, vp: ventral pancreas bud, li: liver, du: duodenum, st:
1.5 Neurog3 is a key transcription factor in endocrine development

The transcription factor Neurog3 has a key function in pancreatic development. As mentioned in the previous sections Neurog3 expression marks the transition of the cells from the pancreas progenitor state to the endocrine precursor state. This is of course an extremely important event that ends the great expansion in the number of cells that forms the pancreatic cell pool and leads to the next phase of fate specification and differentiation of the individual cells. The differentiation process is believed to go through a series of events involving many transcription factors, eventually leading to the formation of mature hormone producing cells organised in the islets of Langerhans, however the process and the sequence of events are not fully understood. The desire to understand the functions of Neurog3 both as a terminator of proliferation and an initiator of differentiation and migration has been the driving force for a substantial amount of research, and the most important findings will be mentioned in the following section.

1.5.1 Basic-helix-loop-helix transcription factors

The Helix-Loop-Helix (HLH) transcriptional regulatory proteins is a family of more than 240 members, that are divided into several classes depending on their structure (Massari and Murre, 2000). The basic-HLH (bHLH) transcription factors are one such class and MyoD, myogenin, and the neurogenins are examples of bHLH transcription factors. The bHLH proteins share the central helix-loop-helix domain with the rest of the HLH family members, but the bHLH transcription factors also contain a basic region adjacent to the HLH region, hence the name. The bHLH transcription factors are involved in developmental events such as lineage commitment and cellular differentiation (reviewed in (Massari and Murre, 2000)). Much knowledge on the neurogenins is based on experiments with other bHLH transcriptin factors e.g. the muscle specific MyoD and myogenin.

The bHLH proteins dimerize with another class of HLH proteins, the E-proteins. The E-proteins E12 and E47 results from different splicing of the E2A gene (reviewed in (Massari and Murre, 2000)). MyoD was shown to dimerize with E47 and induce transcription synergistically in a reporter assay, and MyoD/E47 and MyoD/E12 heterodimers were shown to bind to DNA much stronger than the homodimers of E47, E12, and MyoD (Lassar et al., 1991; Sun and Baltimore,
MyoD homodimer DNA-binding appeared very weak, whereas E12 homodimer DNA-binding was not detected (Sun and Baltimore, 1991). This was attributed to a sequence N-terminally to the basic region, which was believed to inhibit E12 homodimer DNA-binding by not E12 heterodimer DNA-binding which was as efficient as E47 heterodimer DNA-binding (Sun and Baltimore, 1991). The xenopus NGN which is homologue to neurogenin 2 (Neurog2) has been shown to bind DNA in the presence of E12 and Neurog3 was shown to bind DNA in the presence of E47 (Huang et al., 2000; Vosper et al., 2007). The bHLH/E-protein heterodimers bind to E-box elements in the promoters of target genes. The E-boxes consist of the DNA sequence CANNTG, where N can be any base (figure 1.3 C and D) (Ma et al., 1994; Hermann et al., 1998).

The neurogenins are bHLH transcription factors. Neurogenin was first identified in the search for a bHLH factor that would function in determination of the lineage fate of the neurogenic progenitor cells rather than in their differentiation, as neurogenic differentiation 1 (Neurod1) was known to do. That is, it was expected to be expressed at an earlier time point and to regulate Neurod1 expression in neurogenesis, and so neurogenin was identified in Xenopus and mouse (Ma et al., 1996). The search for other bHLH factors that would control earlier stages or different lineages in development of the vertebrate nervous system revealed the identity of two additional neurogenins, neurogenin2 and neurogenin3 (Sommer et al., 1996). They displayed non-overlapping expression patterns in some areas and partial overlapping expression in others, and their expression temporally preceded that of Neurod1 suggesting that the neurogenins would function as determination factors, determining the lineage fate of the cells, and in turn activate the common differentiation factor Neurod1 in the different sublineages (Sommer et al., 1996). It has been shown for Neurog2 that it can direct cell lineage specification and that the choice to mediate this function or induce a general cell fate is determined by phosphorylation of different amino acids; in neural progenitors Neurog2 had a general proneural function, but upon phosphorylation of Y214 (Tyrosine) it induced a dendritic morphology and phosphorylation of S231 and S234 induced motor neuron specification (Hand et al., 2005; Ma et al., 2008).

Neurog3 from mouse was sequenced to be a protein of 214 amino acids with the bHLH domain at amino acid 82-140, however the precise amino acids residues to be included in the bHLH domain varies slightly in different reports (figure 1.3 B) (Ma et al., 1996; Sommer et al., 1996; Watada et al., 2003; Smith et al., 2004). The basic-loop-helix domain is highly conserved between the neurogenins and include a short junction sequence between the basic- and helix1-regions (Sommer et al., 1996). The transcriptional activation domain of Neurog3 was mapped to amino acids 190-214 (figure 1.3 A) (Smith et al., 2004).
Nuclear localisation signals were identified for MyoD, E12, and E47 to be sequences of basic amino acid residues in the basic region, and for the E-proteins also in the N-terminal region (Lingbeck et al., 2003; Lingbeck et al., 2005; Mehmood et al., 2009). It was shown, that the E-protein’s nuclear localisation signals were able to relocate a cytoplasmatic mutant of MyoD to the nucleus quite efficiently, whereas the wild type MyoD was not able to relocate cytoplasmatic mutants of the E-proteins to the nucleus as efficient, rather MyoD nuclear localisation was slightly impaired, suggesting that the nuclear localisation signals of the E-proteins were more important for nuclear localisation of the dimers than the MyoD nuclear localisation signal (Lingbeck et al., 2005). In Neurod1 the basic region acted as a nuclear localisation signal, and in this experiment the nuclear localisation signals of both Neurod1 and E47 were able to locate the dimer to the nucleus when the signal in the other bHLH factor was inactivated (Mehmood et al., 2009). Furthermore, it was shown that the nuclear import of the Neurod1/E47 dimer was energy dependent (Mehmood et al., 2009). As a bHLH factor Neurog3 also contains a region of basic amino acid residues that may act in a similar way as a nuclear localisation signal.

Figure 1.3. The bHLH transcription factors. A. Sketch of Neurog3, indicating a central bHLH domain and a C-terminal activation domain (Sommer et al., 1996; Smith et al., 2004). B. The mouse Neurog3
sequence. Red marks the basic region, green marks the junction region, blue marks helix1, pink marks the loop region, and orange marks helix2 (Sommer et al., 1996). C. Sketch of the MyoD-DNA interaction. MyoD is a bHLH transcription factor and interacts with DNA through the basic region of the bHLH domain. The sketch shows DNA binding of MyoD as a homodimer. Only the bHLH domain is shown. Adapted from (Ma et al., 1994). D. Sketch of the base and phosphate contacts made by a MyoD monomer on DNA. The critical bases CANNTG are indicated with bold lines. The MyoD T115 corresponds to the N89 of Neurog3 (Brennan et al., 1991; Sommer et al., 1996). Adapted from (Ma et al., 1994).

1.5.2 Expression of Neurog3

Neurog3 is expressed in the Pdx1 domain at e.8.5-8.75 and from e.9.0 it can be detected in the dorsal pancreatic bud and from e.9.5 in the ventral. Some of the early Neurog3 expressing cells are detected outside the Pdx1-expressing domain, suggesting that Neurog3 cells can develop independently of Pdx1 expression (Jorgensen et al., 2007). From e11.5 Neurog3 is mainly expressed in the central pancreatic epithelium (Jorgensen et al., 2007). Recently, Neurog3 mRNA was suggested to be expressed in a biphasic pattern at e8.5-e11.0 and from e12 and onwards corresponding to the primary and secondary transitions (Villasenor et al., 2008). This is however highly controversial and not in accordance with the findings of Neurog3 protein in e11.5 pancreas epithelium (Jorgensen et al., 2007). A discrimination of endocrine progenitor cells expressing high versus low levels of Neurog3 was observed in e18.5 mouse embryos (Wang et al., 2009). The low-expressers co-localised with the insulin production marker c-peptide, whereas the high-expressers did not, suggesting a differential function or stage of these subsets of cells (Wang et al., 2009).

Traditionally, Neurog3 expression was believed to cease upon differentiation of the cells, but recently upon reinvestigation of a Neurog3-EGFP reporter mouse low levels of Neurog3 expression was detected in a subset of adult islet cells. This was supported by investigations in two other reporter mouse strains and RT-PCR of adult islets (Wang et al., 2009). Neurog3 was expressed in all islet cell types but only rarely in cells that did not express endocrine hormones (Wang et al., 2009). The expression of Neurog3 in adult islets was suggested to maintain β-cell function (Wang et al., 2009).

Neurog3 expression has been suggested to be induced upon robust pancreas injury in endogenous pancreatic endocrine progenitor cells that do not co-express hormones, and these cells were believed capable of proliferating and differentiating into new β-cells (Xu et al., 2008). The Neurog3 high-expressers identified in e18.5 mouse pancreas were suggested to give rise to the endocrine islet progenitor cells identified in the ducts of adult mouse pancreas (Xu et al., 2008; Wang et al., 2009). However, in another injury model Neurog3 expression was found to be silenced by four endogenous microRNAs, and regeneration of β-cells were
suggested to occur independently of Neurog3 (Joglekar et al., 2007). Hence, the role of putative Neurog3 expression in adult pancreas and regeneration is not quite clear.

1.5.3 Neurog3 function
Neurog3 expression acts as a developmental switch that turns on the endocrine differentiation program (figure 1.4). Studies of genetically modified mice have revealed a lot about Neurog3 function, and as already mentioned mice lacking Neurog3 fail to develop pancreatic endocrine cells, and lineage tracing showed that all types of pancreatic endocrine cells had expressed Neurog3 (Gradwohl et al., 2000; Gu et al., 2002). In mice, ectopic expression of Neurog3 under the pdx-1 promoter induced early differentiation of the pancreatic progenitor cells into glucagon expressing cells rendering the pancreatic buds small as a result of the early termination of the pancreatic progenitor cell expansion (Apelqvist et al., 1999; Schwitzgebel et al., 2000). The same phenotype is seen in mice lacking the Notch1 intracellular mediator delta or the downstream target gene hairy and enhancer of split 1 (Hes1); small pancreatic buds with increased and early expression of Neurog3 and glucagon, again interpreted as premature differentiation of the pancreatic progenitor cells into endocrine cells with a depletion of cells in the pancreatic bud (Apelqvist et al., 1999; Jensen et al., 2000b). The pro-endocrine function of Neurog3 was also demonstrated upon ectopic expression of Neurog3 in the endoderm of the early chicken gut tube (Grapin-Botton et al., 2001). Neurog3 induced migration and differentiation, as the electroporated cells were found in clusters in the mesenchyme, some of them expressing glucagon or somatostatin (Grapin-Botton et al., 2001). These findings supported the hypothesis that Neurog3 expression marks the end of the Notch1 signalling-mediated expansion of the progenitor cell pool and induce endocrine differentiation.
Figure 1.4 Potential functions of Neurog3 in pancreatic development. Neurog3 expression marks the transition from early pancreatic development where the cell compartment grows and expands to late pancreatic development where the exo- and endocrine compartments differentiate and the individual endocrine cell lineages are specified, differentiate, mature and form the islets of Langerhans. At the transition from the early to the late stage in pancreatic development several events are initiated, however it is for the most part unknown if the events are a direct consequence of Neurog3 function, if it is mediated by Neurog3 downstream genes, or if other factors are also necessary for these events to take place.

Recently, the function of Neurog3 as an endocrine switch was challenged, as it was discovered that the cells expressing Neurog3 was not as previously believed irreversibly committed to an endocrine fate (Cras-Meneur et al., 2009). Deletion of presenilins, the catalytic subunit of the Notch1 activator γ-secretase, in Neurog3 expressing progenitor cells directed the cells towards an exocrine fate, suggesting a role of Notch1 in mediating endocrine fate. The mechanism was suggested to be a sequestration of the Notch1 downstream signalling molecule RBPJ from the pancreas specific transcription factor, 1a (Ptf1a, PTF1-p48), that was earlier demonstrated to be necessary for exocrine differentiation (Krapp et al., 1998). Ptf1a and RBPJ were previously shown to co-ordinately induce pdx-1 expression (Miyatsuka et al., 2007; Wiebe et al., 2007). Deletion of the presenilins did not induce exocrine fate in progenitors expressing the transcription factor paired box gene 6 (Pax6), upon deletion of the presenilins,
and the cells at this stage were regarded as being irreversibly committed to the endocrine lineage (Cras-Meneur et al., 2009). This effect of Notch1 was therefore suggested to be involved in the fate choice for only a brief period of time; after Neurog3 was expressed and before Pax6 was expressed (Cras-Meneur et al., 2009). However, the possibility that Neurog3 expressing cells could be directed to an exocrine fate did not change the fact that it is usually observed to induce endocrine fate.

The temporal expression of Neurog3 in development has been shown to influence not only pancreas size and endocrine differentiation, but also the composition of the endocrine compartment, as ectopic Neurog3 expression induced glucagon and somatostatin expressing cells (Apelqvist et al., 1999; Grapin-Botton et al., 2001). This was investigated directly in mice, where Neurog3 was activated at different time points during development in the pdx-1 domain in a Neurog3 null background (Johansson et al., 2007). This resulted in different percentages of the major pancreatic endocrine cell types; early activation of Neurog3 at e8.5 almost exclusively resulted in glucagon expressing cells, whereas later activation of Neurog3 at e14.5 induced a mix of all major cell types with the glucagon expressing cells in a relatively low proportion compared to the proportion of insulin or PP expressing cells (Johansson et al., 2007). This effect of changing the onset time of Neurog3 expression might be a result of a longer-lasting Neurog3 signal in the cells where Neurog3 was expressed early, and hence a longer period of time for different cell types to develop. However, this was not the full explanation, as Neurog3 expression from e10.5-e14.5 and e14.5-e18.5; and from e12.5-e14.5 and 14.5-16.5 gave the cells four days or two days respectively to differentiate but at different time points of pancreas development. The fact that different cell type distributions were observed for the same Neurog3 exposure time suggested that the competency of the progenitors would change over time perhaps as a result of differential expression of other transcription factors depending on the embryonic day (Grapin-Botton et al., 2001).

The finding that the endocrine output induced by Neurog3 expression may vary depending on the presence of other transcription factors were investigated at another level in recent in vivo studies. Neurog3 was co-expressed together with other pro-differentiation transcription factors in adult animals mediated by viral infection and this showed amazing results with regard to reprogramming of adult cells to insulin producing cells. In one experiment streptozotocin-induced diabetic mice were treated with adenovirus mediated expression of Neurog3-v-maf musculoaponeurotic fibrosarcoma oncogene family protein A (Mafa) or Neurog3-Mafa-Neurod1 in the liver. This induced insulin and glucagon production, reduced hyperglycaemia five days after infection, and restored glucose tolerance as determined by an intraperitoneal glucose tolerance test (Song et al., 2007). In another setup, adenoviral infection of the pancreas of
adult mice with Pdx-1, Neurog3, and Mafa showed that it was possible to reprogram adult pancreatic exocrine cells into cells that resembled β-cells; they expressed insulin and several β-cell markers, induced local angiogenic remodelling and ameliorated hyperglycaemia in diabetic mice in up to nine weeks. Pdx-1 and Mafa expression was sustained after two months even though transgene expression was undetectable and Neurog3 expression had ceased, suggesting that the effect was stable (Zhou et al., 2008). These experiments show the potential of the right transcription factor cocktail to induce adult cell reprogramming and pose new possibilities for treatment of e.g. diabetes that would be independent of large-scale stem cell to β-cell differentiation facilities.

As previously mentioned, it was recently discovered that Neurog3 was expressed in adult pancreatic islets cells (Wang et al., 2009). The function of Neurog3 in differentiated embryonic and adult tissue was investigated by inactivating Neurog3 in insulin-producing cells at embryonic stages and in pdx-1 expressing endocrine cells in adult animals. This lead to impaired endocrine function in mature β-cells and mature islet cells as measured by marked increase in fasting blood glucose values, and reduced expression of Neurog3 downstream genes (Neurod1, paired box gene 4 (Pax4), and myelin transcription factor 1 (Myt1)) and β-cell markers (Mafa and insulin) (Wang et al., 2009). Thus, the presence of Neurog3 was required not only for initiation of endocrine differentiation in pancreas development but also for promotion of islet cell maturation, maintenance and function (Wang et al., 2009).

In humans, two Neurog3 mutants (R107S and R93L) have been identified. These mutations were the cause of congenital malabsorptive diarrhea with an almost absence of intestinal enteroendocrine cells. The patients were born non-diabetic, but the patients with the Neurog3R93L mutation developed type 1 diabetes at the age of 8, while the patient with the Neurog3R107S mutation died at the age of two (Wang et al., 2006). The mutations occurred in the bHLH region of Neurog3 and were shown to impaired Neurog3 activity in vitro (Jensen et al., 2007). The effect of these mutations seemed to be more severe with regard to intestinal function than for endocrine pancreas function.

### 1.5.4 Regulation of Neurog3

Neurog3 could potentially be regulated in several ways, e.g. at the level of expression. Several transcription factors interact directly with the Neurog3 promoter to activate transcription (figure 1.5).
Figure 1.5. Regulation of Neurog3 and its downstream targets. The figure shows a simplified model of a selection of the signals regulating Neurog3 expression and the Neurog3 downstream targets. The upper grey area marks the early pancreatic development where the progenitor cells proliferate and the cell pool expands. The lower grey area marks the late pancreatic development where the cell types are specified and later mature into hormone producing cells. Red boxes represent Neurog3 inhibitory signals, green boxes represent Neurog3 inducing signals, and blue circles represent Neurog3 downstream signals. 1: (Apelqvist et al., 1999); 2: (Hald et al., 2003); 3: (Jensen et al., 2000b); 4: (Lee et al., 2001); 5: (Hart et al., 2003); 6: (Norgaard et al., 2003); 7: (Harmon et al., 2004); 8: (Lynn et al., 2007); 9: (Jacquemin et al., 2000); 10: (Oliver-Krasinski et al., 2009); 11: (Smith et al., 2003); 12: (Smith et al., 2004); 13: (Watada et al., 2003); 14: (Mellitzer et al., 2006); 15: (Breslin et al., 2002); 16: (Marsich et al., 2003); 17: (Huang et al., 2000); 18: (Anderson et al., 2009); 19: (Wang et al., 2008).

One cut domain, family member 1 (Onecut1, HNF6) bound to the Neurog3 promoter and induced transcription (Jacquemin et al., 2000). Mice lacking Onecut1 had at e13.5 and e14.5 developed virtually no Neurog3 expressing cells, and later only a few Neurog3 expressing cells that did not migrate from the ductal epithelium were observed (Maestro et al., 2003). Prolonged Onecut1 expression under the pdx-1 promoter resulted in diabetic animals with hyperplastic islets containing an increased number of α-, δ-, and PP-cells, and dysfunctional β-cells with absent or severely reduced expression of the GLUT-2 glucose transporter (Gannon et al., 2000). However, if this phenotype also involved a prolonged Onecut1 mediated Neurog3 expression is unknown.
Pdx-1 induced transcription from the Neurog3 promoter and transcription was increased cooperatively in the presence of Onecut1 (Oliver-Krasinski et al., 2009). It was also shown that the transcription factors forkhead box A2 (Foxa2, HNF3β) and HNF1 homeobox A (Hnf1a) bound to the Neurog3 promoter (Lee et al., 2001). SRY-box containing gene 9 (Sox9) binds to the promoter of Neurog3 and induce transcription (Lynn et al., 2007). Sox9 also bound to the promoters of Foxa2, Onecut1, and Sox9 itself (Lynn et al., 2007). Sox9, Foxa2, Hnf1a, and are involved in maintenance of the pancreatic progenitor cell pool, in differentiation of α–cells, and in the normal function of β–cells, respectively (Pontoglio et al., 1998; Lee et al., 2005; Seymour et al., 2007). The Neurog3 target Myt1 has been shown to be potentiate Neurog3 expression in a feed-forward loop, whereas Neurog3 itself binds to its own promoter and represses transcription suggesting a negative feed-back loop (Smith et al., 2004; Wang et al., 2008). The transcription factors that induce Neurog3 expression seem to work in a complex network (figure 1.5).

Notch1 was shown to inhibit Neurog3 mediated differentiation of in ovo electroporated endodermal cells into glucagon and somatostatin expressing cells and to prevent Neurog3-mediated expression of several downstream genes (Ahnfelt-Ronne et al., 2007a). This effect of Notch1 may act through its downstream target Hes1 which was shown to bind to the Neurog3 promoter with the function of a repressor and block Neurog3 transcription (Lee et al., 2001). Whether the Notch1-mediated repression of Neurog3 is direct or indirect, it may serve to maintain the undifferentiated state of the pancreatic progenitor cells allowing proliferation and expansion of the cell pool before differentiation begins. Hence, the balance of Hes1 and Notch1 signaling on one side and the Neurog3 inducing transcription factors expressed in the pancreas area on the other side may be involved in regulation of Neurog3 expression and the resulting endocrine differentiation. How exactly the progenitor cells escape Notch1 signaling is not known although the mechanism of lateral inhibition has been suggested. According to this theory, an equilibrium of Notch1 receptors and ligands exists on adjacent cells, and when the equilibrium is skewed towards more ligands on one cell, differentiation is initiated in that cell whereas the adjacent cell remains undifferentiated. However, what might skew such equilibrium and if this mechanism is functioning in the developing pancreas is unknown.

The mesenchyme surrounding the developing pancreas has been reported to play a role in the expansion of the progenitor cell pool and also on Neurog3 expression, although the mechanisms are not so well defined. The absence of mesenchyme in e13.5 rat pancreas culture induced Neurog3 expression and insulin production which was not seen in co-cultures with rat mesenchyme (Duvillie et al., 2006). The repressive effect on pancreas epithelial Neurog3 expression required cell-cell contact and was a result of Notch1-signalling. Forced
expression of fibroblast growth factor 10 (Fgf10) was also shown to inhibit Neurog3 expression through Notch1 signalling (Hart et al., 2003; Norgaard et al., 2003). Latent transforming growth factor beta binding protein 1 (Ltbp1, TGFβ) signalling was also suggested to play a role in pancreas development, as mice lacking the Ltbp1 family ligand Gdf11 developed smaller pancreatic buds with increased Neurog3 expression and premature endocrine differentiation (Harmon et al., 2004).

Histone deacetylases are involved in chromatine condensation and decrease the accessibility of the transcription factors to target genes. Neurog3 mRNA levels in rat pancreas explants increased upon inhibition of class I and II histone deacetylases, and recently Neurog3 mRNA was suggested to be more widespread in the pancreatic epithelium than Neurog3 protein (Haumaitre et al., 2008; Villasenor et al., 2008). This might imply a regulation of the amount of Neurog3 at the post-transcriptional level, and the findings that Neurog3 expression could be silenced by endogenous microRNAs supported this, although the mechanism and impact of such post-transcriptional regulation is unknown (Joglekar et al., 2007).

Post-translational modifications and protein-protein interactions have also been suggested to play a role in the regulation of Neurog3. In the Xenopus NGN, which is a homologue of Neurog2, phosphorylation of the amino acid T118 has been shown to stabilize the protein and increase the half life from approximately 20 to 60 minutes (Vosper et al., 2007). Perhaps related mechanisms apply to Neurog3. The Akt kinase has been shown to increase Neurog3 transcriptional activity in the presence of the transcriptional activator E1A binding protein p300 (Ep300) and to strengthen the interaction between these two proteins (Vojtek et al., 2003). Neurog3 binds directly to Foxa2 and Hnf1a through the bHLH domain, and the combination of Neurog3 with these transcription factors mediate transcription of the target genes Pax4 and NK2 transcription factor related, locus 2 (Nkx2-2) (Smith et al., 2003; Watada et al., 2003).

1.5.4 Neurog3 downstream targets

Neurog3 mediates transcription of several genes that are implicated in the differentiation and fate-specification of endocrine progenitors towards mature hormone producing cells (figure 1.5).

Mice lacking Neurog3 did not express Neurod1, and analysis of the Neurod1 promoter revealed several E-boxes as putative Neurog3 binding sites (Gradwohl et al., 2000; Huang et al., 2000). Neurog3 in combination with E47 was demonstrated to bind to E-box elements in the Neurod1 promoter and activate transcription in luciferase assays (Huang et al., 2000). The combination
of Neurog3 and E47 induced a larger response in the reporter than observed with any of the proteins alone (Huang et al., 2000). In adult reporter mice Neurod1 was expressed in islets of Langerhans, in the developing pancreas it partially co-localised with Neurog3, and injection of Neurog3 in xenopus embryos induced exogenous Neurod1 expression (Huang et al., 2000). A linear correlation was observed for a higher level of Neurog3 expression and a higher level of Neurod1 expression in mouse pancreatic adenocarcinoma cells and in adult pancreatic ductal epithelial cells (Boretti and Gooch, 2007). Recently, Nkx2-2 was shown to bind to the Neurod1 promoter in an electrophoresis mobility shift assay and mediate transcription in luciferase assays (Anderson et al., 2009). Furthermore, Neurod1 expression in the developing pancreas of mice lacking Nkx2-2 was markedly decreased, and luciferase assays showed a cooperative effect of Neurog3 and Nkx2-2 in Neurod1 expression (Anderson et al., 2009). It was suggested that Neurog3 and Nkx2-2 form a large transcriptional complex on the Neurod1 promoter that would tightly regulate the transcriptional output (Anderson et al., 2009).

Mice lacking Neurog3 did not express Pax4, and promoter analysis of Pax4 revealed the presence of an E-box, a potential target sequences for Neurog3 (Gradwohl et al., 2000; Smith et al., 2003). Neurog3 and its binding partner E47 did not mediate transcription of Pax4 alone but did so when co-transfected with E47 and either Hnf1a or HNF1 homeobox B (Hnf1b) and especially Hnf1a was potent in this respect (Smith et al., 2003). Deletion analysis of Neurog3 showed that in combination with E47 and Hnf1a full length Neurog3 and Neurog376-214 induced transcription from the Pax4 minienhancer, whereas Neurog31-131 and Neurog376-131 were unable to do so (Smith et al., 2003). Hnf1a interacted directly with Neurog3 within the region of amino acid 76-214 (Smith et al., 2003). The binding of Neurog3 to a Pax4 promoter E-box and transcriptional activity on the promoters were later confirmed in another study where Hnf1a did not seem to be required for Neurog3 mediated transcription (Smith et al., 2004). As observed for Neurod1, a higher level of Neurog3 expression mediated by retrovirus transfection lead to a higher level of Pax4 target gene expression in mouse pancreatic adenocarcinoma cells and to a lesser extent in adult pancreatic ductal epithelial cells (Boretti and Gooch, 2007).

Analysis of the Nkx2-2 promoter revealed the presence of several E-boxes that would be potential targets for Neurog3 (Watada et al., 2003). Binding sites for Foxa2 were also detected and both Neurog3/E47 and Foxa2 were shown to bind to sequences in the promoter (Watada et al., 2003). Neurog3/E47 induced transcription from the promoter whereas Foxa2 did not, but co-transfection with the two transcription factors showed a synergistic effect on transcription of Nkx2-2 (Watada et al., 2003). Deletion studies of Neurog3 showed that together with E47 the full length Neurog3, Neurog31-131, and Neurog376-214 induced low levels
of transcription that were enhanced 2-3 fold by the presence of Foxa2, whereas the bHLH domain, interpreted in this study as Neurog3\textsuperscript{76-131}, did neither induce transcription alone or in the presence of Foxa2, suggesting that either the C- or N-terminal end of Neurog3 was required for this function (Watada et al., 2003). Furthermore, it was shown that Neurog3 and Foxa2 interacted directly and the interaction was mapped to their DNA-binding domains, for Neurog3 that is the bHLH domain (amino acid 76-131 in this report) and for Foxa2 that is a winged helix domain (Watada et al., 2003). The binding of Neurog3 to a Nkx2-2 E-box and transcriptional activity on the promoters were later confirmed in another study (Smith et al., 2004).

Neurog3 was also shown to bind to the rat insulin I promoter and E-box elements from the Neurog3 promoter itself (Smith et al., 2004). However, the binding of Neurog3 to its own promoter did not induce transcription, rather it was repressed, and when E47 was present this was even more pronounced (Smith et al., 2004). As no repressor domain was known in Neurog3 the repressive effect was speculated to be the result of competition with an unidentified activating factors for binding of the E-box (Smith et al., 2004).

Neurog3 was able to bind to the promoter of insulinoma-associated 1 (Insm1, IA1) and mediate transcription in a luciferase assay (Mellitzer et al., 2006). Neurod1 displayed the same abilities, but Insm1 was present in the embryonic pancreas of mice lacking Neurod1 whereas it was not detected in mice lacking Neurog3, suggesting that NeuroD1 was not required for Neurog3-induced Insm1 transcription (Mellitzer et al., 2006).

Neurog3 was shown to bind to the atonal homolog 8 (Atoh8, also known as Math6) promoter and induce transcription in a luciferase assay, and at e15.5 Atoh8 is expressed in almost 50% of the Neurog3 expressing cells in the developing mouse pancreas (Lynn et al., 2008). However, Atoh8 expression was not restricted to the endocrine compartment and is probably also expressed by Neurog3 independent routes and the role of Math8 in pancreas development remains unclear (Lynn et al., 2008).

Luciferase assays with a Myt1 reporter plasmid showed that Neurog3 was able to induce transcription of Myt1 and mice lacking Neurog3 showed reduced Myt1 expression (Wang et al., 2008). However, the fact that Myt1 was not completely absent in these mice suggested that Myt1 expression was not exclusively dependent on Neurog3 (Wang et al., 2008).
1.5.5 The function of Neurog3 target genes and other selected transcription factors

How the lineage specification of the endocrine cells is regulated has been the subject of many speculations. It seems likely that the insulin- and glucagon-producing cells differentiate as two independent cell lineages with a common Neurog3 expressing precursor, as a lineage tracing study showed that the mature α- and β-cells did not stain for the same genes to be active during development, and an expression profile study suggested that a precursor cell expressing Neurog3, Pdx-1, and NK6 homeobox 1 (Nkx6-1) would split into two lineages with different gene expression profiles (Herrera, 2000; Jensen et al., 2000a). A complex network of transcription factors is activated throughout endocrine differentiation, and the differential expression or regulation of these may account for the specification of the endocrine cell types. Often the generation of transgenic mice carrying a mutation in the gene of interest has been used to characterise the function of the transcription factors, and the role of the Neurog3 downstream genes in endocrine differentiation will be mentioned in this section (figure 1.6).

![Figure 1.6. Selected transcription factors in late pancreatic development.](image)

The figure is a simplified model of a selection of the transcription factors involved in late pancreatic development. The stages the
transcription factors (green boxes) function at are marked by arrows and are based on studies of mice lacking the specific transcription factor. This is a simplified model as the transcription factors probably function on several stages and in combination with other factors expressed intra- or extracellular, since they are expressed for longer periods of time and in some instances even in mature hormone producing cells. Circles represent the stage of differentiation of the cells and arrows between them marks a series of undefined events leading to the next stage in differentiation. Broken lines mark indirect evidence that Arx and Pax4 function as repressors in the differentiation of the specific cell type. Please note in the text that in mice lacking both Arx and Pax4 the number of somatostatin expressing cells was increased (Collombat et al., 2005). 1: (Sosa-Pineda et al., 1997); 2: (Collombat et al., 2003); 3: (Matsuoka et al., 2004); 4: (Gradwohl et al., 2000); 5: (Gu et al., 2002); 6: (Schonhoff et al., 2004); 7: (St-Onge et al., 1997); 8: (Sander et al., 1997); 9: (Naya et al., 1997); 10: (Sussel et al., 1998); 11: (Sander et al., 2000); 12: (Prado et al., 2004); 13: (Gu et al., 2004); 14: (Wang et al., 2007); 15: (Sander et al., 2000); 16: (Artner et al., 2006); 17: (Artner et al., 2007).

Mice lacking Neurod1 had severe diabetes and died a few days after birth (Naya et al., 1997). They had abnormal islet morphogenesis with small clusters of endocrine cells expressing multiple hormones, a reduced number of endocrine cells, and in particular the number of β-cells was affected (Naya et al., 1997). Neurod1 activates the Pax6 promoter together with the transcriptional co-factor E47, and the lack of Pax6 expression in Neurog3 -/- mice may be a result of the loss of Neurod1 although it is not known if Neurog3 is able to induce Pax6 directly (Gradwohl et al., 2000; Marsich et al., 2003).

Pax4 was detected in the fetal pancreas at e10.5 and onwards peaking around e13.5, later it was present in the β-cells only, and it was absent in adult islets (Sosa-Pineda et al., 1997; Smith et al., 1999). Mice lacking Pax4 do not develop mature insulin- and somatostatin-producing islet cells, but have a considerably higher numbers of glucagon producing cells, suggesting a role for Pax4 in specification of β- and δ-cells (Sosa-Pineda et al., 1997). ε-cells were also reported to replace β-cells in mice lacking pax4 (Prado et al., 2004). Based on promoter and reporter studies Pax4 was suggested to be a transcriptional repressor that would suppress α-cell differentiation and permit β-cell differentiation during pancreas development (Smith et al., 1999). Ectopic expression of Pax4 in the pdx-1 and Pax6 domains in mice produced oversized islets primarily consisting of insulin expressing cells, and when expressed in mature α-cells converted these into insulin expressing cells (Collombat et al., 2009).

Nkx2-2 is expressed in the dorsal pancreas bud from e8.75 and in the ventral bud from e9.5 (Jorgensen et al., 2007). Mice lacking Nkx2-2 showed a disrupted islet structure, reduced numbers of α- and PP-cells, and a lack of mature β-cells although β-cell precursors were detected (Sussel et al., 1998; Sander et al., 2000). ε-cells were reported to replace β-cells in mice lacking Nkx2-2, a phenotype similar to what was observed for mice lacking Pax4, suggesting that ε- and β-cells share a common progenitor, and that Nkx2-2 and pax4 are required to specify or maintain differentiation of the β-cell fate (Prado et al., 2004).
Inhibition of Insm1 in e12.5 mouse dorsal pancreas explants by morpholinos showed a marked reduction in insulin and glucagon expression compared to the control after four days in culture whereas amylase expression was unaffected (Mellitzer et al., 2006). Insm1 was identified to be a transcriptional repressor, that could potentially inhibit expression of Neurod1 (Breslin et al., 2002).

Myt1 is expressed at low levels in the pancreas area at e8.5, at e13.5 it was detected in a scattered fashion in the pancreas bud, and at e15.5 expression was much reduced (Gu et al., 2004). Myt1 expression partially overlapped with Neurog3 expression in e10.5-e15.5 mice (Wang et al., 2007). A dominant negative form of Myt1 was expressed in endocrine progenitors under the Neurog3 promoter, and this inhibited endocrine differentiation as determined by a reduction in the number of cells expressing glucagon and insulin at e 14.5 (Gu et al., 2004). Mice lacking Myt1 had abnormal adult islet cells co-expressing multiple hormones, reduced Glut2 expression, attenuated glucose induced insulin secretion, and died postnatally (Wang et al., 2007). These experiments suggested a role of Myt1 in endocrine differentiation that was supported by the findings that Myt1 was able to partially rescue Neurog3 from Notch1 mediated inhibition in in ovo electroporated early gut tube chicken endoderm (Ahnfelt-Ronne et al., 2007a). Myt1 is also described in section 1.5.6.

Other genes that are not regulated by Neurog3 are known to play a role in endocrine differentiation, and the most established ones are mentioned in the following (figure 1.6).

Nkx6-1 is expressed early in the ventral bud around e8.75-e9.0 and again from e10.5. From e9.0 it is also expressed in the dorsal bud. Later it becomes restricted to the central epithelium, and this is opposite the pattern of Ptf1a which was seen in the growing tips (Jorgensen et al., 2007). Mice lacking Nkx6-1 fail to develop β-cell precursors and mature β-cells (Sander et al., 2000). The Nkx6-1/Nkx2-2 double mutant mouse showed the same phenotype as the Nkx2-2 mutant mouse suggesting that Nkx2-2 functions upstream of Nkx6-1 in β-cell development, although the loss of Nkx6-1 was more severe with the loss of β-cell progenitors (Sander et al., 2000).

Pax6 is expressed in the pancreatic epithelium a little later than e9.0 and at e11.5 the Pax6 expressing cells begin to cluster (Jorgensen et al., 2007). Mice homozygous for a mutant allele of Pax6, believed to be a null allele, had a reduction in number of all endocrine cell types, whereas mice lacking Pax6 were reported especially to fail to develop glucagon producing cells, although the number of insulin producing cells was also dramatically reduced (St-Onge et al.,
These investigations suggest a role of Pax6 in endocrine differentiation and perhaps especially in α-cell differentiation.

Aristateless related homeobox (Arx) was detected in the pancreas area at e9.5, later in the epithelium, in the differentiating precursors, and in the forming islets (Collombat et al., 2003). Mice lacking Arx did not develop mature α-cells and had increased numbers of β- and δ-cells, which was the opposite than what had been observed for Pax4 (Collombat et al., 2003). The fact that the lack of both Arx and Pax4 changed the distribution of hormone producing cells in opposite directions was the basis of the idea that they together determine the fate of the progenitor cells to become either α-cells or β-cells. The lack of both Arx and Pax4 resulted in the loss of both α- or β-cells and an increase in the number of δ-cells supporting this theory (Collombat et al., 2005).

Mafa is present in insulin producing cells in the developing pancreas from e13.5, and was also co-expressed with insulin in mature β-cells, but was never detected in cells producing glucagon or somatostatin (Matsuoka et al., 2004). Mafa was able to induce insulin production in the non-insulin producing cell line αTC6, supporting a role of Mafa in β-cell specification (Matsuoka et al., 2004).

v-maf musculoaponeurotic fibrosarcoma oncogene family protein B (Mafb) is expressed in the developing pancreas in both insulin and glucagon producing cells and in proliferating cells not expressing hormones, but in the adult islets it is restricted to α-cells (Artner et al., 2006). Mice lacking Mafb had a reduced number of insulin and glucagon expressing cells e10.5-e15.5 whereas the total number of endocrine cells was unchanged at e.12.5-e.18.5, suggesting a role of Mafb in the development of both α- and β-cells (Artner et al., 2007).

1.5.6 The Neurog3 downstream target Myt1
Myt1 is one of three vertebrate CHC-type zinc-finger transcription factors that include Myt1, MyT1L (Png1), MyT3 (Nzf3, St18). Only My11 is expressed in the embryonic pancreas (Wang et al., 2007). Myt1 was first identified in the nervous system as a transcription factor of the zinc finger family containing six zinc finger motifs (Kim and Hudson, 1992). In xenopus, Myt1 was shown to be induced by X-NGNR-1 and negatively regulated by Notch1 signalling, and mutation studies showed that it had a function in differentiation of oligodendrocyte progenitors but not in astrocyte progenitors, suggesting a role in lineage determination (Bellefroid et al., 1996; Nielsen et al., 2004). Furthermore, Notch1 inhibited X-NGNR-1 mediated expression of Myt1 (Bellefroid et al., 1996). Myt1 was discovered in the pancreas in an expression analysis.
and was found to exist in two forms, one with six zinc fingers and one with seven zinc fingers (figure 1.7) (Matsushita et al., 2002; Gu et al., 2004; Wang et al., 2008). The seven zinc finger form was expressed at higher levels than the six zinc finger form especially in newly generated neurons (Matsushita et al., 2002). The Myt1 expression pattern and what is known about its function in pancreatic development was described in section 1.5.5. The precise function of Myt1 in endocrine differentiation and if there is a difference in the function of the two forms of Myt1 are still unknown.

**Figure 1.7. Myt1.** A. The Myt1 promoter contains transcription start sites, one for transcription of the Myt1 7 zinc finger form and one for transcription of the Myt1 6 zinc finger form. B. The Myt1 7 zinc finger form contains 7 zinc finger motifs. C. The Myt1 6 zinc finger form contains 6 zinc finger motifs. Modified from (Nielsen et al., 2004).
2. Objectives

The overall objective of this thesis was to understand in greater detail the role of the transcription factor Neurog3 in the differentiation of pancreatic progenitor cells towards an endocrine fate.

As already mentioned, mice lacking Neurog3 fail to develop the entire endocrine compartment of the pancreas and lineage tracing has shown that all pancreatic endocrine cells had at some point expressed Neurog3 (Gradwohl et al., 2000; Gu et al., 2002; Schonhoff et al., 2004). Several other transcription factors are involved in endocrine differentiation, but investigations of genetically modified mice lacking these transcription factors revealed phenotypes that suggested that they operate later in development than Neurog3. That is they didn't lack the entire endocrine compartment but displayed abnormalities that suggested that the event occurred after the endocrine fate had been specified, such as abnormal islet morphology, impaired maturation of the islet cells, reduced numbers of endocrine cells, or the lack of some but not all types of hormone producing cells. Some of these genes were shown to be regulated by Neurog3 (figure 1.5 (Naya et al., 1997; Sander et al., 1997; Sosa-Pineda et al., 1997; St-Onge et al., 1997; Sussel et al., 1998; Gradwohl et al., 2000; Sander et al., 2000; Gu et al., 2002; Collombat et al., 2003; Gu et al., 2004; Prado et al., 2004; Schonhoff et al., 2004; Artner et al., 2006; Artner et al., 2007; Wang et al., 2007)).

It is not known exactly how Neurog3 specifies endocrine cell fate. With this report we try to enlighten the role of Neurog3 as the inducer of endocrine differentiation by addressing this question from different angles. It was previously reported from deletion studies and reporter assays that the bHLH domain was not sufficient to induce transcription from a reporter and the activation domain was mapped to the C-terminal end in in vitro reporter studies, but does the endocrine specifying function of Neurog3 map to certain protein domains in vivo (Smith et al., 2003; Watada et al., 2003)? Neurog3 is believed to function as a heterodimer with one of the E-proteins like other bHLH proteins do (Sun and Baltimore, 1991; Massari and Murre, 2000). Would co-expression of E-proteins with Neurog3 in over-expression studies enhance the endocrine differentiation function of Neurog3? Mutations were discovered in Neurog3 in human patients with severe congenital malabsorptive diarrhea (Wang et al., 2006). How would these mutations affect the function of Neurog3 as the inducer of endocrine differentiation in the pancreas? At the time that this project was initiated Myt1 was recently identified as a novel transcription factor involved in endocrine differentiation (Gu et al., 2004). Is Myt1 a direct target of Neurog3? With this thesis we try to answer these questions in order to enlighten the
immediate responses to Neurog3 expression and how Neurog3 initiates endocrine differentiation (figure 2.1).

**Figure 2.1. How does Neurog3 induce endocrine differentiation in pancreatic development.**
Many questions remain to be answered with regard to how Neurog3 is capable of initiating endocrine differentiation. The figure shows the questions that will be addressed in the following sections.

### 2.1 Neurog3 protein domains

A more detailed understanding of Neurog3 function would shed light to how pancreatic progenitor cells are directed into an endocrine fate, and we wanted to map protein domains in Neurog3 that would be important for its function in vivo. Deletion analysis of Neurog3 has previously been used to identify regions required for protein-protein interactions and transcriptional activity on different promoters using Gal4 fusions (Smith et al., 2003; Watada et al., 2003; Smith et al., 2004). These studies had revealed that the bHLH domain was not sufficient to induce transcription in vitro from reporters containing the Pax4 and Nkx2-2 promoters and that the C-terminal end of Neurog3 contained the activation domain required to induce expression from the Neurod1 promoter. However, the effects of these findings were not
investigated *in vivo*. We wanted to use a similar deletion analysis strategy to characterize Neurog3 in more detail and identify protein domains required for its endocrine inducing function *in vivo*. To investigate this we would test truncated Neurog3 proteins for their cellular localisation, for their ability to bind to DNA, to induce transcription from the Neurod1 promoter, and to induce endocrine differentiation by *in ovo* electroporation of the early chicken gut tube endoderm.

### 2.2 The role of E12 in Neurog3 function

The E-proteins E12 and E47 are generated by alternative splicing of the E2A gene, and as mentioned these two proteins function as heterodimerization partners for bHLH proteins (Sun and Baltimore, 1991; Massari and Murre, 2000). We speculated that the observations achieved with the truncated Neurog3 proteins might be affected by a potential lack of sufficient amounts of endogenous dimerization partners in the cell lines and chicken embryo cells, as transfection and electroporation are believed to result in a substantial over expression of the gene of interest. We therefore wanted to investigate if co-expression of a dimerization partner would enhance the effects observed with the truncated Neurog3 proteins and perhaps compensate for the observed loss of function of some of the Neurog3 proteins.

E47 homodimers binds to DNA and potentially has effects on its own in cells whereas E12 homodimers are believed not to bind DNA and to be inactive (Sun and Baltimore, 1991; Massari and Murre, 2000). MyoD/E12 and MyoD/E47 heterodimers were previously shown to bind DNA with equally efficient, and we chose E12 as the dimerization partner in our studies, as we wanted to avoid potential effects of the E47 homodimer (Sun and Baltimore, 1991; Massari and Murre, 2000).

### 2.3 Neurog3 mutants in human

Two naturally occurring mutations of Neurog3 were discovered in humans (Wang et al., 2006). The mutations had detrimental effects on enteroendocrine differentiation and caused severe congenital malabsorptive diarrhea, whereas the pancreatic endocrine function was not affected at birth, and only later in childhood was diabetes manifest (Wang et al., 2006). The mechanism behind this difference was not explained, and we wanted to investigate the nature of these human Neurog3 mutants and the effects that the mutations would have on Neurog3
function and activity, in order to explain the presence of functional pancreatic endocrine cells in the absence of enteroendocrine cells.

2.4 Neurog3 as the activator of Myt1 transcription

As earlier mentioned, Myt1 was discovered in the pancreas in an expression analysis, and expression of a dominant negative form of Myt1 under the Neurog3 promoter inhibited endocrine differentiation and lead to a reduction in the number of cells expressing insulin and glucagon at e14.5 in mice (Gu et al., 2004). In the nervous system Myt1 had been demonstrated to be involved in lineage specification of the neuronal subtypes of cells, and in xenopus it was induced by X-NGNR-1 and inhibited by Notch1 signalling (Bellefroid et al., 1996; Nielsen et al., 2004). Myt1 was discovered in two forms containing either six or seven zinc fingers (Matsushita et al., 2002). We wanted to investigate if Myt1 was a direct target of Neurog3.
3. Neurog3 protein domains

The structure and function of a protein are closely related features. Knowing the structure of a protein gives clues to its function and the way it is regulated. The amino acid sequence and the overall structure of Neurog3 as a bHLH protein and its function as a transcription factor is well known and the activation domain was mapped by others (Sommer et al., 1996; Smith et al., 2004). It would be interesting to characterise Neurog3 further and map features of Neurog3 that would be essential for its function as a transcription factor, e.g. the region necessary for nuclear localisation, and regions of importance for its transcriptional activity or in vivo effects.

In order to map such features, we used a deletion assay approach. If the deleted regions would contain any essential regulatory elements the function of the specific truncated protein would be lost or impaired. On the basis of such findings it would be possible to investigate that specific region in detail and perhaps to reveal the identity of the amino acids responsible. The truncated Neurog3 proteins were characterised for expression on western blots using lysates from transfected cells, DNA binding on electrophoresis mobility shift assays (EMSA), cellular localisation in transfected cells, transcriptional activity using luciferase assays, and in vivo function using in ovo electroporation. Some of the results concerning the in vitro and in vivo activity of the truncated Neurog3 proteins entered into manuscript 1 (included as appendix 1), and so did supporting data such as the expression studies (western blot) and the cellular localisation studies (immunofluorescence stainings of transfected cells).

3.1 Mutation strategy

Regions of approximately 30, 60, and 90 amino acids were deleted from each end of the protein resulting in six truncated Neurog3 proteins (Neurog3\textsuperscript{1-178}, Neurog3\textsuperscript{1-150}, Neurog3\textsuperscript{1-124}, Neurog3\textsuperscript{29-214}, Neurog3\textsuperscript{56-214}, Neurog3\textsuperscript{91-214}). Furthermore, truncated proteins consisting of the bHLH domain alone, the bHLH domain plus the N-terminal region, or the bHLH domain plus the C-terminal region was generated (Neurog3\textsuperscript{74-138}, Neurog3\textsuperscript{1-138}, and Neurog3\textsuperscript{74-214}). Neurog3 with a point mutation in the basic region was also designed to serve as a negative control (Neurog3\textsuperscript{N89D}). The N89D mutation was chosen on the basis of results from the bHLH transcription factor myogenin (Myog) that showed that a T to D point mutation in the corresponding amino acid (T87) disabled Myog from DNA binding while retaining its ability to bind to E12 (Brennan et al., 1991). This amino acid is highly conserved as either T or N in bHLH proteins and the corresponding amino acid in MyoD T115 was also identified as DNA binding (Brennan et al., 1991; Ma et al., 1994). N89 in Neurog3 was a good candidate for such
mutation and we anticipated that the N89D mutation would have the same effect as the T87D mutation in Myog. The truncated Neurog3 proteins are presented in figure 3.1. The cloning of the Neurog3 constructs is described in detail in section 7.1 and was also described in the manuscript enclosed as appendix 1.

![Figure 3.1. Schematic presentation of the truncated Neurog3 proteins. All constructs contain an N-terminally located HA-epitope tag for easy detection.](image)

### 3.2 Expression of the truncated Neurog3 proteins in cells

#### Purpose
In order to draw any conclusions on the truncated Neurog3 proteins in further studies it was essential to know if they were at all expressed in cells or if the mutations were detrimental.

#### Results
3T3 cells were transfected with the Neurog3 constructs. Two days later transfection efficiency was checked and the cells were harvested. The protein content was measured and the lysates adjusted accordingly before they were subjected to SDS-PAGE and western blotting. The membranes were incubated with antibodies against the HA-epitope present in all Neurog3 constructs and EGFP as a control of the amount of plasmid in the loaded lysates. The
procedure is described in detail in section 7.6. All lysates contained approximately the same amount of EGFP although the band from the Neurog3\textsuperscript{374-138} sample was a bit weaker and the bands from the Neurog3\textsuperscript{31-124} and Neurog3\textsuperscript{91-214} samples were a bit stronger (figure 3.2). Neurog3 and Neurog3\textsuperscript{N89D} gives a band of approximately 30 kDa and the truncated Neurog3 proteins gives bands of lower molecular weight (figure 3.2). The Neurog3 constructs with C-terminal truncations gave strong bands, but the ones with N-terminal truncations are very weak, and the bands from Neurog3\textsuperscript{74-214}, Neurog3\textsuperscript{91-214} and Neurog3\textsuperscript{374-138} are barely detectable (figure 3.2). Another primary antibody was tested giving the same result (data not shown).

![Figure 3.2. Expression of the Neurog3 constructs in transfected 3T3 cells. Lysates were normalised with regard to protein content before loading on the gel.](image)

**Figure 3.2. Expression of the Neurog3 constructs in transfected 3T3 cells.** Lysates were normalised with regard to protein content before loading on the gel.

**Conclusions**

The obvious explanation to why the Neurog3 construct with N-terminal truncations were not detectable to the same degree as the other Neurog3 constructs would be that they were expressed at a much lower level than the other Neurog3 constructs or that they were degraded at a higher rate. In immunofluorescence stained cells these constructs are detectable but at a weaker intensity (figure 3.3). However, as described in later sections the Neurog3 constructs with N-terminal truncations also showed very weak bands when generated *in vitro* (figure 3.5), suggesting that the weak bands were not a result of degradation or instability of the protein in the cells. The most striking observation in this regard was that the *in vitro* activity of these truncated proteins was at the same level as Neurog3 (figure 3.8). Taken together, the conclusion must be that all Neurog3 constructs are expressed but that some are harder than others to detect with antibodies. This may be a result of the HA-tag being located at the N-
terminus. The immediate surroundings of the tag changes whenever a new region is deleted from the N-terminal end of Neurog3. If this interferes with antibody binding resulting bands on the membrane would be weaker. In order to avoid this, a short linker region could have been introduced in the design process to divide the tag and the gene. The western blot was also commented in the manuscript enclosed as appendix 1.

3.3 Cellular localization of the truncated Neurog3 proteins

Purpose

Transcription factors are normally located in the nucleus, as this is necessary for a transcription factor to exert its function. If some of the truncated proteins are not located in the nucleus it will suggest a nuclear localisation signal, or other motifs essential for nuclear localisation has been removed.

Results

3T3 and P19 cells were transfected with the Neurog3 constructs. Two days later the cells were fixed and immunofluorescence stained with antibodies against the HA-tag to visualise the truncated proteins and against EGFP to stain transfected cells. DAPI stain was also applied to stain all nuclei. Pictures were obtained at a confocal microscope. The procedure is described in section 7.8. Pictures obtained from the P19 cells were not as illustrative as the pictures obtained from the 3T3 cells, as the P19 cells have a much smaller cytoplasm (data not shown). Wild type Neurog3 was located in the nucleus and so was Neurog3N89D and most of the truncated proteins (Neurog31-178, Neurog31-150, Neurog31-138, Neurog329-214, and Neurog356-214; figure 3.3 B, C, D, E, G, H, and I). However, the Neurog3 mutants that had a truncated bHLH domain, the bHLH domain alone, and the truncated proteins that had lost the N-terminal domain were also present to a larger degree in the cytoplasm (Neurog31-124, Neurog374-214, Neurog391-214, and Neurog374-138; figure 3.3 F, J, K, and L). In general, the proteins with N-terminal truncations stained weaker than Neurog3 and the other truncated Neurog3 proteins, and this was most pronounced for Neurog374-214, Neurog391-214, and Neurog374-138. At western blots these mutants were also very weak (figure 3.2). These results were also mentioned in the manuscript enclosed as appendix 1.
Figure 3.3. Cellular localisation of the truncated Neurog3 proteins. 3T3 cells were transfected with the truncated Neurog3 proteins and immunofluorescence stained. Red marks the HA-tagged Neurog3 constructs, green marks EGFP in transfected cells, and grey marks all nuclei. The panels show representative sections of pictures obtained on a confocal microscope using the 25x oil objective.

Conclusions
The N89D point mutation did not interfere with nuclear localisation. An intact bHLH domain was required for Neurog3 to be located primarily in the nucleus, however, when the N- and C-terminal ends were lost the intact bHLH alone was present to a larger degree in the cytoplasm. The loss of the C-terminal domain did not interfere with nuclear localisation whereas the loss of the N-terminal domain markedly increased the presence of Neurog3 in the cytoplasm. The amino acids flanking the bHLH domain on the N-terminal side must be important for nuclear localisation (within the region of amino acid 57-73). One reason for this could be that these amino acids might encode or be part of a nuclear localisation signal which corresponds well with the fact that the basic region of other bHLH factors has been shown to encode a nuclear localisation signals (Lingbeck et al., 2003; Mehmood et al., 2009). Alternatively, they might be important for proper folding of the protein or protein-protein interactions e.g. with E12. These are both factors that could potentially influence nuclear localisation.

3.4 Cellular localisation of the truncated Neurog3 proteins in the presence of wild type Neurog3

Purpose
The cellular localisation of the truncated Neurog3 proteins were investigated in section 3.3. To investigate the observed differences in cellular localisation further, the truncated Neurog3 proteins were co-transfected with wild type Neurog3 into cells. If the localisation of the truncated proteins and the wild type proteins would differ, the truncations would have affected amino acids involved in nuclear localisation. If the cellular localisation would overlap, the truncations would not have affected amino acids involved in nuclear localisation.

Results
3T3 cells were co-transfected with myc-tagged Neurog3 and the HA-tagged truncated Neurog3 proteins. Two days post transfection the cells were fixated and subjected to immunofluorescence staining and confocal microscopy as described in section 7.8. Neurog3 wild type and most of the truncated Neurog3 proteins with an intact bHLH domain were located primarily in the nucleus. Neurog3\textsuperscript{91-214} and Neurog3\textsuperscript{74-138} were shown to be present in the cytoplasm to a larger degree (figure 3.3 K and L) and that was also the case when they were co-transfected with wild type Neurog3 (figure 3.4 K and L), and to some extent that was true also for Neurog3\textsuperscript{74-214} (figure 3.4 J). Neurog3\textsuperscript{1-124} was also present to a larger degree in the cytoplasm in figure 3.3, but when co-transfected with wild type Neurog3 this seems to be less pronounced and both the cytoplasm and nucleus seem to stain to some degree for both proteins (figure 3.4 F).

\textbf{Figure 3.4. Cellular localisation of the truncated Neurog3 proteins in the presence of wild type Neurog3.} 3T3 cells were co-transfected with the truncated Neurog3 proteins and myc-Neurog3 and immunofluorescence stained. Blue marks the HA-tagged Neurog3 constructs, red marks myc-tagged wild type-Neurog3, green marks EGFP in transfected cells, and grey marks all nuclei. The panels show representative sections of pictures obtained on a confocal microscope using the 25x oil objective.
Conclusions

Wild type Neurog3 and most of the truncated proteins were located in the nucleus. Neurog3^{74-214}, Neurog3^{91-214} and Neurog3^{74-138} were present to a larger degree in the cytoplasm although the wild type protein located primarily to the nucleus, suggesting that the region affected by the truncations may be involved in nuclear localisation (74-91). The fact that the bHLH domain alone and the Neurog3 protein that lacks the N-terminal end are located to a larger degree in the cytoplasm suggests that the region flanking the bHLH domain on the N-terminal side was also involved in nuclear localisation as found in section 3.3. The presence in the cytoplasm of Neurog3^{1-124} was less pronounced when it was co-transfected with wild type Neurog3, and wild type Neurog3 was less restricted to the nucleus, than when these proteins were transfected independently. The mechanism behind this is unknown.

3.5 DNA binding of the truncated Neurog3 proteins

Purpose

The ability of a transcription factor to bind to the DNA of its target gene promoter is essential for its function. If some of the truncated Neurog3 proteins would be unable to bind DNA the deleted region must be important for this feature of Neurog3.

Results

The ability of the truncated Neurog3 proteins to bind to DNA was tested using an electrophoresis mobility shift assay (EMSA). The EMSA was performed with protein generated in vitro. The procedure was optimised and control experiments performed to identify the bands as described in section 7.4 and figure 7.2. As a control of plasmid expression the in vitro generated protein was subjected to SDS-PAGE and western blotting (figure 3.5). It was not possible to measure the protein content of the in vitro generated protein samples directly because of the red colour of the TNT® Quick Master Mix, so a ponceau staining of the membrane and a coomassie staining of a parallel gel was performed to evaluate the loading of the samples on the gel (figure 3.5 B and C). The size of the bands in the separate lanes of the ponceau and coomassie stains were approximately the same meaning that the protein content of the samples were also approximately the same (figure 3.5 B and C). The western blot showed weaker bands for the proteins with N-terminal deletions. This was also what was seen on the western blot of the Neurog3 protein generated in cells (figure 3.2).
Neurog3 proteins generated in the absence of E12 did not bind to the DNA-probe at a detectable level, except for Neurog3\textsuperscript{74-214} which showed a very faint band (figure 3.6 A). Neurog3 generated in the presence of E12 did bind the DNA-probe and so did Neurog3\textsuperscript{1-178}, Neurog3\textsuperscript{1-150}, Neurog3\textsuperscript{29-214}, Neurog3\textsuperscript{56-214}, Neurog3\textsuperscript{74-214}, and Neurog3\textsuperscript{74-138} (figure 3.6 B). The following Neurog3 proteins generated in the presence of E12 did not bind the DNA-probe or did so at undetectable levels: Neurog3\textsuperscript{1-138}, Neurog3\textsuperscript{1-124}, Neurog3\textsuperscript{91-214}, and Neurog3\textsuperscript{N89D}. Neurog3\textsuperscript{74-214} generated with E12 showed a black smear on the EMSA and the sample had to be diluted 10 times in order to give a distinct band (figure 3.6 B and C). These results are also commented in the manuscript enclosed as appendix 1.
Figure 3.6. DNA binding of the truncated Neurog3 proteins. In vitro generated truncated Neurog3 proteins were subjected to EMSA in order to investigate their ability to bind to DNA. A. EMSA of the truncated Neurog3 proteins generated in vitro. B. EMSA of the truncated Neurog3 proteins generated in vitro in the presence of E12. * Sample diluted 10 times. C. Section of an EMSA of the truncated Neurog3 proteins generated in vitro in the presence of E12 in which the Neurog3^{74-214} sample was not diluted 10 times.

It was also tested if potential Neurog3 homodimers were able to bind to DNA and the truncated Neurog3 proteins were generated in vitro in the presence of myc-tagged Neurog3. Neurog3 was not able to bind to the DNA probe as a homodimer at a detectable level and neither was any of the truncated Neurog3 proteins generated in the presence of wild type Neurog3 (figure 3.7).
Figure 3.7. DNA binding of the truncated Neurog3 proteins in the presence of wild type Neurog3. Neurog3 did not bind to the DNA probe as a homodimer at a detectable level, and neither did any of the Neurog3 constructs when generated in vitro in the presence of wild type myc-Neurog3.

Conclusion

DNA binding of Neurog3 was greatly increased in the presence of E12 and without E12 DNA binding was absent or too weak to be detected at this EMSA. The fact that Neurog3N89D did not bind DNA or did so at an undetectable level, supported the idea behind the design of this construct, which was that this construct would encode a non-DNA-binding Neurog3 mutant that could be used as a negative control. The truncated Neurog3 proteins with an affected bHLH domain did not bind DNA or did so at an undetectable level, and this was also the case for Neurog31-138 which had lost the C-terminal end. Neurog324-214 which had lost its N-terminal end showed extensive DNA-binding estimated to approximately 10 times the level of Neurog3 DNA-binding. This is remarkable considering that this protein was hardly detectable on western blots and that the staining of this protein in cells was quite faint. This supports the speculation that antibody detection of the N-terminally located HA-tag was affected in the Neurog3 proteins with N-terminal deletions.
3.6 *In vitro* activity of the truncated Neurog3 proteins

**Purpose**

In order to map any putative regions of Neurog3 necessary for transcription the truncated Neurog3 proteins were characterized with regard to *in vitro* activity using luciferase assays.

**Results**

The reporter consisted of the Neurod1 promoter driving the firefly luciferase gene and the activators were expression plasmids encoding Neurog3 or the truncated Neurog3 proteins. As a control for transfection efficiency a renilla luciferase-expression plasmid was co-transfected into the cells. For further information on the assay and plasmids see section 7.9 where optimisation of the luciferase assay is also described. We tested the luciferase assay in both 3T3 cells and P19 cells, but the 3T3 cells gave very poor results. For some reason the three fold induction by Neurog3 that was seen in the first test assays (described in section 7.9) was not reproducible in the 3T3 cells (figure 3.8 A and figure 7.3). In P19 cells in contrast the Neurod1 reporter showed a very good response to Neurog3 of around 25 in average.

Experiments in P19 cells show that when the bHLH domain was affected by truncations or the N89D point mutation the Neurog3 mutants were not able to induce a response on the Neurod1 promoter (Neurog3\(^{1-124}\), Neurog3\(^{91-214}\), Neurog3\(^{N89D}\), \(p<0.05\), figure 3.8 B). The bHLH domain alone and the bHLH domain plus the N-terminal end were not able to induce a response either (Neurog3\(^{74-138}\), Neurog3\(^{1-138}\), \(p<0.05\), figure 3.8 B). Deleting several to all amino acids C-terminally to the bHLH domain did not influence transcription activity (Neurog3\(^{29-214}\), Neurog3\(^{36-214}\), Neurog3\(^{74-214}\), \(0<0.05\), figure 3.8 B). From the C-terminal end the outermost 36 amino acids could be deleted without any effect on transcriptional activity (Neurog3\(^{1-178}\), \(p<0.05\), figure 3.8 B), whereas deleting the 64 most C-terminal amino acids reduces transcriptional activity to around 50% (Neurog3\(^{1-150}\), \(p<0.05\), figure 3.8 B).
**Figure 3.8. In vitro activity of the truncated Neurog3 proteins on the Neurod1 promoter.**
Luciferase assays were performed on the Neurod1 promoter. The values represent the response of the activators in relation to the basal activity of the reporter (fold induction). A. In 3T3 cells the truncated Neurog3 proteins gave very poor results and the induction level of Neurog3 of around three that was seen in the optimization experiments could not be reproduced. B. In P19 cells Neurog3 induced a response from the reporter of around 25 fold in average. n= 4-9, bars represent standard error of the mean, * marks $p < 0.05$ compared to the basal activity level of the reporter, dots mark $p < 0.05$ compared to Neurog3.

Conclusions
The bHLH domain alone bound to DNA but was not sufficient to induce transcription (figure 3.6 and 3.8 B). The N-terminal domain did not contain any regulatory elements that were essential for transcriptional activity on the Neurod1 promoter as all of this could be lost without any effect on the transcriptional activity of the mutant compared to Neurog3. Deleting the most terminal amino acids from the C-terminal end did not have any effect on the transcriptional activity, but deleting 64 amino acids reduced the transcriptional activity to approximately 50% and deleting the entire C-terminal domain abolished transcription. The most C-terminal amino acids 190-214 have previously been reported to hold the activation domain of Neurog3 (Smith et al., 2004). The results shown here do not support this finding. However, the more central region of the C-terminal end of Neurog3 was important for transcriptional activity, as truncating this region impaired transcriptional activity. As described above, this was not a result of poorer DNA binding (section 3.5), or poorer nuclear localization (section 3.3). This assay was also described in the manuscript enclosed as appendix 1.
3.7 *In vivo* activity of the truncated Neurog3 proteins

**Purpose**

To identify regions of Neurog3 required for its function and test the significance of the reduced levels of activity displayed by some of the truncated proteins in the luciferase assay, the truncated proteins were analysed for their ability to induce migration and differentiation as evaluated by glucagon expression *in vivo*.

**Results**

The *in vivo* activity of the truncated Neurog3 proteins was examined using *in ovo* electroporation and whole mount immunofluorescent stainings and confocal microscopy of the embryos. The details are described in the section 7.10-7.12.

Neurog3 induced migration in a large fraction of the electroporated cells as seen by the GFP expressing cells located away from the gut tube (figure 3.9 B). Neurog3 also induced differentiation in some of the electroporated cells as seen by the GFP expressing cells co-expressing glucagon (yellow cells, figure 3.9 B). Cells electroporated with the empty vector expressed GFP and the electroporated cells stay in the gut tube (figure 3.9 A). Neurog3\textsuperscript{N89D} was considered inactive as electroporated cells stayed in the early gut tube and failed to induce glucagon expression (figure 3.9 C). This was also the case for Neurog3\textsuperscript{1-124} and Neurog3\textsuperscript{91-214} in which the bHLH domain was truncated. The bHLH domain alone (Neurog3\textsuperscript{74-138}) was also unable to induce differentiation and migration (figure 3.9 D). The Neurog3 construct with N-terminal truncations and an intact bHLH domain were all able to induce migration and differentiation, and even the one that had lost the entire N-terminal domain was active *in vivo* (figure 3.9 I, J, and K). The Neurog3 constructs with C-terminal deletions were also active, although the protein that had lost the entire C-terminal domain only induced glucagon expression and migration in two out of six embryos. These results were also presented in the manuscript enclosed as appendix 1.
Figure 3.9. *In vivo* activity of the truncated Neurog3 proteins. Pictures show stacks of optical sections projected into one picture of a representative embryo electroporated with the truncated Neurog3 protein. Parenthesis state the number of embryos with the shown phenotype observed out of the total number of embryos electroporated with the truncated Neurog3 protein. Red marks glucagon, green marks EGFP in the electroporated cells, and grey marks Foxa2 in the early gut tube. Pictures were obtained with a 10x objective.

The above mentioned results were based on visual appearance of the electroporated chicken embryos. In order to see if the ability of the active Neurog3 mutants to induce differentiation varied the number of electroporated cells (expressing GFP) and the number of electroporated cells that had differentiated (co-expressing GFP and glucagon) was counted on optical sections of the confocal stacks as described in section 7.13. Figure 3.10 shows representative optical
sections from chicken embryos electroporated with the Neurog3 constructs and figure 3.11 shows the results of the cell counting. There was no difference between the active Neurog3 constructs in their ability to induce differentiation. This was also mentioned in the manuscript enclosed as appendix 1.

**Figure 3.10. In vivo activity of the truncated Neurog3 proteins.** The pictures show a representative optical sections from a confocal stack of a whole mount immunofluorescence stained chicken embryo electroporated with the Neurog3 constructs. Pictures were obtained with a 10x or 20x objective.
Figure 3.11. Quantification of differentiation. *In ovo* electroporated chicken embryos were subjected to whole mount immunofluorescence staining for glucagon and EGFP and pictures were obtained with cofocal microscopy. Optical sections from the electroporated embryos were evaluated for the fraction of GFP expressing cells co-expressing glucagon. A student’s t-test was applied to determine significance. Bars represent standard error of the mean, n= 2-5.

Judged by visual appearance there may be a tendency for the lower activity Neurog3 mutants (Neurog3¹-¹⁵₀ and Neurog3¹-¹³₈, figure 3.9 F and G) to be less efficient in mediating migration than the higher activity Neurog3 mutants (figure 3.9 E, I, J, and K).

Conclusions
The truncated Neurog3 proteins with an intact bHLH domain all induced differentiation of early gut tube endodermal cells into glucagon expressing cells, however, the bHLH domain alone was not sufficient to induce differentiation. The fact that Neurog3¹-¹³₈ only induced differentiation in two out of six of the embryos suggested that Neurog3 activity needs to be above a certain threshold for differentiation and migration to occur and that this protein has an activity level close to that threshold. This would allow for other variables such as minor variations in the temperature of the incubator or the electroporation efficiency of each embryo to determine whether the electroporated cells differentiate or not. The active Neurog3 mutants did not differ in their ability to induce differentiation as determined by the fraction of electroporated cells co-expressing glucagon.
3.8 Neurog3 mediated hormone expression

Purpose
The truncated Neurog3 proteins with \textit{in vivo} activity all induced glucagon to the same degree. In our laboratory we usually do not see insulin expression in Neurog3 electroporated chicken embryos. We speculated that Neurog3 might have a role in fate specification and that the truncated proteins might reveal any regions of importance for this function.

Results
Immunofluorescense stainings for insulin were performed on whole mounted chicken embryos electroporated with the truncated Neurog3 proteins with the largest truncations that still showed activity (Neurog3_{1-138} and Neurog3_{74-214}, figure 3.12). Neurog3_{1-138} did not induce insulin expression in the endoderm of the early chicken gut tube, but Neurog3_{74-214} did in three out of seven embryos (figure 3.12). This was also described in the manuscript enclosed as appendix 1.

Figure 3.12. Induction of insulin expression by two of the truncated Neurog3 poteins. Insulin stainings of \textit{in ovo} electroporated chicken embryos subjected to whole mount immunofluorescence staining. Top panel pictures show confocal stacks from an embryo projected into one picture and lower panel pictures show single optical sections. The confocal stacks were obtained with a 10x objective and the optical sections with a 20x objective. Arrows mark insulin staining in clusters of GFP expressing cells and the box marks the area enlarged in (C').
Conclusions
Electroporation of the early chicken gut tube endoderm with Neurog3 usually does not induce any hormones other than glucagon. In the manuscript enclosed as appendix 1 it was described that a longer incubation time of the embryos after electroporation would allow for PP and somatostatin to be expressed in Neurog3 electroporated cells, but insulin expression was never induced. It was surprising that Neurog3^{74-214} was able to induce insulin expression two days after electroporation in the duodenum, and it may suggest that Neurog3 may be able to induce endocrine lineage specification rather than only a general endocrine fate. It would be highly interesting to support this finding with further investigations, and examine the other truncated proteins for induction of insulin expression with the purpose of mapping the responsible regions and explaining the mechanism.

Manuscript 1.
The transcriptional activation potential of Neurog3 affects migration and differentiation of ectopic endocrine cells in the chicken endoderm.

In manuscript 1 (appendix 1) we addressed the question if the activity level and temporal expression of ectopic Neurog3 would affect migration and differentiation of endocrine cells in the chicken endoderm. The manuscript was a result of the collaboration of several people employed at the Hagedorn Research Institute. I did the experiments regarding the activity level of Neurog3 presented in Figure 1, 2, 4, and supplementary figures 1 and 2. These results are also presented in the above section. The experiments presented in figure 3 were performed by Dr. Hani Yassin, and the experiments presented in figure 5, 6, and 7 and supplementary figure 3 were performed by Merete L. Lafon, M.Sc. The first drafting of the manuscript was done by me supervised by Dr. Jacob Hecksher-Sorensen, and Dr. Jan Nygaard Jensen participated in scientific discussions and evaluation of the text. The manuscript is planned to be submitted to the journal Developmental Dynamics.
4. The role of E12 in Neurog3 function

Background
The E-proteins E12 and E47 were previously identified as Neurog3 binding partners and are present in many tissues (reviewed in (Massari and Murre, 2000)). In section 3.5 of this thesis it was shown that E12 was necessary for Neurog3 DNA binding. The cell- and chicken-based assays with Neurog3 and the truncated Neurog3 proteins relied on the endogenous E-proteins of the cell lines or the chicken embryos to complex with Neurog3. One could speculate that in these experiments there has been a great excess of Neurog3 compared to E-proteins, and that some of the findings described in the previous section could be a result of an insufficient amount of endogenous E12 to dimerize with the huge amount of exogenous Neurog3. In order to investigate this, some of the experiments described in the previous section were repeated with the modification that exogenous E12 was added and these experiments will be described in this section. We tested the cellular localisation to see if the introduction of exogenous E12 would rescue the truncated Neurog3 proteins that were present in a larger degree in the cytoplasm to the nucleus, the in vivo activity of the truncated Neurog3 proteins in the presence of exogenous E12 to see if the inactive proteins could be rescued, and did preliminary experiments on the in vitro activity of the truncated Neurog3 proteins in the presence of exogenous E12. In the assays where in vitro generated Neurog3 was used, E12 was already included and will not be dealt with in the following (section 3.5). When examining the effect of E12 on Neurog3 activity it would have been highly relevant to knock down endogenous E12 in cells and chicken embryos and perform the studies under this condition as well.

4.1 The role of E12 in the cellular localisation of Neurog3

Purpose
It would be relevant investigate if the presence of exogenous E12 would influence the localisation of Neurog3 or the truncated Neurog3 proteins. This was of course primarily relevant for the truncated Neurog3 proteins that were present to a larger extent in the cytoplasm and not exclusively in the nucleus (figure 3.3 F, J, K, and L).

Results
3T3 cells were co-transfected with E12 and Neurog3 or the truncated Neurog3 proteins and pictures were obtained with confocal microscopy as described in section 7.8. Neurog3 and
most of the truncated Neurog3 proteins were primarily located in the nucleus in the presence of exogenous E12, as was the case in the absence of exogenous E12 (figure 4.1 B, C, D, E, G, H, and I and figure 3.3 B, C, D, E, G, H, and I). Neurog3\(^{91-214}\) and Neurog3\(^{74-138}\) were still present to a larger extent in the cytoplasm after exogenous E12 was introduced (figure 4.1 K and L), and Neurog3\(^{74-214}\) too was present in the cytoplasm to some extent (figure 4.1 J). Neurog3\(^{1-124}\) was present to a larger degree in the cytoplasm when transfected alone (figure 3.3 F), but when exogenous E12 was added it located primarily to the nucleus (figure 4.1 F).

![Figure 4.1. Cellular localisation of the truncated Neurog3 proteins in the presence of exogenous E12.](image)

3T3 cells were co-transfected with the truncated Neurog3 proteins and E12 and immunofluorescence stained. Red marks the HA-tagged Neurog3 constructs, green marks EGFP in transfected cells, and grey marks all nuclei. The panels show representative sections of pictures obtained on a confocal microscope using the 25x objective.

Conclusions

Neurog3 and most of the truncated Neurog3 proteins were located primarily in the nucleus regardless of whether exogenous E12 was added or not. The supply of exogenous E12 changed the localisation of Neurog3\(^{1-124}\) from extensive presence in the cytoplasm to primarily being located in the nucleus. Neurog3\(^{91-214}\), Neurog3\(^{74-138}\), and to some extent Neurog3\(^{74-214}\) are still present to a larger extent in the cytoplasm. The fact that exogenous E12 influences the localisation of the protein with C-terminal truncation of the bHLH domain more than the proteins with an intact or N-terminally truncated bHLH domain and no N-terminal end suggested that the amino acids flanking the bHLH domain on the N-terminal side were more important in terms of nuclear localisation or dimerization with E12 than the amino acids in the C-terminal end of the bHLH domain or C-terminally to it. This corresponds well with previous
findings that the basic region is responsible for nuclear localisation in other bHLH factors (Lingbeck et al., 2003; Mehmood et al., 2009). However, it is not known if the basic region alone is responsible for nuclear localisation, as putative E12-based motifs might be important for nuclear localisation too. This is of course pure speculation. Knockdown of endogenous E12 would show if Neurog3 would be able to locate to the nucleus by the basic region alone, but unfortunately it was not possible to investigate this within the time course of this project. Neurog3\textsuperscript{1-124} has a truncated loop2 region which would be expected to interfere with E12 dimerization. When transfected alone this protein was clearly present in the cytoplasm but co-transfection with E12 relocated it to the nucleus. This may reflect that the endogenous E12 was titrated away, or perhaps a lower affinity between the two proteins that is overcome by the over-expression of both. The fact that Neurog3\textsuperscript{74-214}, Neurog3\textsuperscript{91-214} and Neurog3\textsuperscript{74-138} were present to a larger degree in the cytoplasm even in the presence of exogenous E12 supports the findings from the localisation studies (section 3.3 and 3.4) that suggest that this is a result of a directly impaired nuclear localisation signal.

4.2 In vivo activity of Neurog3 in the presence of exogenous E12

Purpose
It would be relevant to investigate if the presence of exogenous E12 would influence the differentiation and migration inducing potential of Neurog3 or the truncated Neurog3 proteins. It was of course primarily interesting to investigate if the activity of the truncated Neurog3 proteins that were inactive in figure 3.9 C, D, H, and L could be rescued by addition of E12.

Results
Chicken embryos were \textit{in ovo} co-electroporated with E12 and Neurog3 or the truncated Neurog3 proteins. The embryos were subjected to wholemount preparation, immunofluorescence staining, and confocal microscopy as described in section 7.10-7.12. Electroporating chicken embryos with E12 alone generated some clustering of the electroporated cells that may either be a result of altered adhesion or proliferation properties of the E12 over-expressing cell (figure 4.2 A and 4.3 A). This fact made it harder to interpret the pictures obtained for some of the truncated Neurog3 proteins. Neurog3 co-electroporated with E12 induced massive migration of the electroporated cells but practically no differentiation or clustering of the cells (figure 4.2 B and 4.3 B). This was unexpected on the basis of the findings from electroporations with Neurog3 alone where migration, clustering and differentiation were observed (figure 3.9 B). The proteins that were identified as active in figure 3.9 showed the same effects as Neurog3 and E12 in this study also, that is migration
but no differentiation or clustering (figure 4.2 E, F, I, J, and K and figure 4.3 E, F, I, J, and K). The proteins that were identified as inactive in figure 3.9 were also inactive when co-electroporated with E12, although as mentioned it was a bit hard to determine in some instances as E12 alone mediated some clustering of the electroporated cells although the cells remained in the endoderm (figure 4.2 C, D, H, and L and figure 4.3 C, D, H, and L). Neurog3\textsuperscript{1-138} that was able to mediate differentiation and clustering in 2/6 of the embryos when electroporated alone but did not show any obvious activity when co-electroporated with E12. However, it is possible that a potential weak activity would go unnoticed because the effects might be masked by the E12 mediated effects.

<table>
<thead>
<tr>
<th>E12</th>
<th>Neurog3</th>
<th>Neurog3\textsuperscript{N89D}</th>
<th>Neurog3\textsuperscript{74-138}</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="A" alt="Image" /></td>
<td><img src="B" alt="Image" /></td>
<td><img src="C" alt="Image" /></td>
<td><img src="D" alt="Image" /></td>
</tr>
<tr>
<td>(6/6)</td>
<td>(5/5)</td>
<td>(3/3)</td>
<td>(3/3)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Glucagon/EGFP/FoxA2</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="E" alt="Image" /></td>
</tr>
<tr>
<td>Neurog3\textsuperscript{1-178}</td>
</tr>
<tr>
<td>(7/7)</td>
</tr>
</tbody>
</table>

| ![Image](I) | ![Image](J) | ![Image](K) | ![Image](L) |
| Neurog3\textsuperscript{29-214} | Neurog3\textsuperscript{56-214} | Neurog3\textsuperscript{74-214} | Neurog3\textsuperscript{91-214} |
| (8/8) | (11/11) | (7/7) | (5/5) |
Figure 4.2. *In vivo* activity of the truncated Neurog3 proteins in the presence of exogenous E12. Pictures show confocal stacks of optical sections projected into one picture of a representative embryo co-electroporated with E12 and Neurog3 or the truncated Neurog3 proteins. Parenthesis state the number of embryos with the shown phenotype observed out of the total embryos electroporated with the respective truncated Neurog3 protein. Red marks glucagon, green marks EGFP in the electroporated cells, and blue marks Foxa2 in the early gut tube. Pictures were obtained on a confocal microscope using a 10x objective.

![Figure 4.2](image1.png)

Figure 4.3. *In vivo* activity of the truncated Neurog3 proteins in the presence of exogenous E12. The pictures show a representative optical sections from a confocal stack of a whole mount immunofluorescence stained chicken embryo co-electroporated with E12 and Neurog3 or the truncated Neurog3 proteins. Pictures were obtained on a confocal microscope using a 10x or 20x objective.

![Figure 4.3](image2.png)

**Conclusions**

Co-electroporation with E12 did not rescue any of the inactive truncated proteins to become active. Neurog3<sup>1-138</sup> showed some activity in two out of six embryos when electroporated alone but no activity in six out of six embryos when co-electroporated with E12. However, this may be a result of the E12-mediated effects on the cells masking potential subtle effects induced by Neurog3<sup>31-138</sup>, or perhaps of two few repetitions of the experiment. The fact that E12 was able to relocate Neurog3<sup>31-124</sup> to the nucleus in 3T3 cells did not result in any *in vivo* activity. On the
basis of results from the EMSA that showed no DNA binding of this mutant, this was expected (figure 3.6 B). The fact that adding exogenous E12 to this assay abolished differentiation and clustering of the cells electroporated with Neurog3 is puzzling since Neurog3-E-protein interaction is usually regarded necessary for Neurog3 activity (reviewed in (Massari and Murre, 2000)). E12 did not inhibit the migration induced by Neurog3, rather, it tended to enhance this process. This supports the idea that the differentiation and delamination effects induced by Neurog3 are distinct processes (Ahnfelt-Ronne et al., 2007a). The mechanisms by which these processes might operate as uncoupled are unknown. However, it has been suggested that a higher Neurog3 activity would result in less differentiation and more migration whereas a low activity would result in more differentiation and less migration (Manuscript 1, appendix 1). If this was the case it would be relevant to suggest that adding exogenous E12 might increase activity of Neurog3 and favour migration over differentiation of the electroporated cells. This could be examined in a luciferase assay (section 4.3). E12 was shown to increase xenopus-NGN, and if this would apply to Neurog3 this might also increase activity (Vosper et al., 2007). An increased Neurog3 activity might also be a consequence of an increased half-life. However, the effect of exogenous E12 on the Neurog3 response could also be indirect, as E12 clearly had some effects independent of Neurog3 possibly mediated by other transcription factors or binding partners endogenously present in the cells.

4.3 In vitro activity of Neurog3 in the presence of exogenous E12

Purpose
When chicken embryos were co-electroporated with Neurog3 and E12 the electroporated cells did not differentiate but migrated extensively. It was previously shown that the E-proteins enhanced Neurog3 activity (Huang et al., 2000). Perhaps the increased migration and reduced differentiation observed in section 4.2 was a result of an E12-mediated increase in Neurog3 activity.

Results
In order to investigate if E12 would enhance Neurog3 activity a luciferase assay was performed. P19 cells were transfected with the Neurod1 reporter, Neurog3 (60 or 20 ng per well), and E12 (20 or 60 ng per well), the cells were checked for transfection efficiency two days later and harvested for analysis as described in section 7.9. E12 alone (20 ng) did not have any transcriptional activity on the Neurod1 reporter (figure 4.4). Neurog3 alone (60 ng) induced a response of approximately 20-fold compared to the basal level of the reporter (figure 4.4). Addition of E12 (20 ng) did not increase the activity of Neurog3, and surprisingly,
and a slight decrease was observed, although this was not significant (figure 4.4). When Neurog3 was co-transfected with a surplus of E12 (20 ng Neurog3 and 60 ng E12) this also tended to decrease the activity of Neurog3 compared to the situation where cells were transfected with equal amount of Neurog3 and E12 (20 ng of each) (figure 4.4). Again this was not significant.

**Figure 4.4. In vitro activity of Neurog3 co-transfected with E12.** Luciferase assays were performed with the Neurod1 promoter, Neurog3 (20 or 60 ng per well), and E12 (20 or 60 ng per well). The values represent the response of the activators in relation to the basal activity of the reporter (fold induction). Bars represent standard error of the mean, * marks p<0.05 compared to the basal activity level of the reporter, lines indicate no statistical difference between the two values, n=8-9.

**Conclusions**

Addition of E12 to the luciferase assay did not increase the activity of Neurog3. Rather a tendency of E12 to reduce Neurog3 activity was observed although this was not significant. This was unexpected as E47 was previously shown to enhance the transcriptional activity of Neurog3 on E-box elements of the Neurod1 promoter (Huang et al., 2000). The endogenous level of E-proteins must have been sufficient for Neurog3 dimerization and activity, and the tendency of exogenous E12 to reduce the Neurog3 response may be a result of sequestering of general transcription factors from the Neurog3/E12 complexes by E12 or E12-complexes with proteins other than Neurog3. The mechanism by which exogenous E12 favors Neurog3 mediated migration over differentiation in the chicken electroporation study (figure 3.9) could
not be explained by an increase or decrease in Neurog3 activity. E12-mediated effects may inhibit differentiation either directly or indirectly. Studies from our lab suggested that cells co-electroporated with Neurog3 and E12 were cycling, and maybe this would prevent differentiation of the cells (Yassin, 2008). The cycling of cells would explain the clumps of EGFP expressing cells observed in chicken embryos electroporated with E12 alone, and maybe the cycling would somehow inhibit the pro-differentiation signals mediated by Neurog3. This is however pure speculation.
5. Studies of the Neurog3 mutants discovered in humans

Neurog3 mutants in human were first described in 2006 (Wang et al., 2006). The patients were identified shortly after birth with congenital malabsorptive diarrhea. This was due to a lack of or very few enteroendocrine cells in the gut tube. The patients had a normal pancreas function from birth but two out of three developed diabetes at the age of eight, while the third patient died before that age. The Neurog3 mutants were identified as Neurog3<sub>R93L</sub> and Neurog3<sub>R107S</sub>. We used the functional assay described in the above sections to investigate the function of these mutants and give a comment on how these mutants would be able to induce endocrine differentiation in the pancreas when it was impaired in the gut tube.

The results were described in the publication enclosed as appendix 2. The cloning of Neurog3<sub>R93L</sub>, Neurog3<sub>R107S</sub>, and Neurog3<sub>R93L,R107S</sub> was performed by Dr. Jan Nygaard Jensen, who also did the luciferase assays, chicken electroporations with Neurog3<sub>R93L</sub>, Neurog3<sub>R107S</sub>, Neurog3<sub>R93S,R107S</sub> mutants, and the writing of the article. I developed the assay, did chicken electroporations with the controls (EGFP, Neurog3, and Neurog3<sub>N89D</sub>, pictures A, B, and C in the published article enclosed as appendix 2) and did the supporting experiments described in the following; EMSA and in ovo co-electroporations with E12.

5.1 DNA binding of the human Neurog3 mutants

Purpose
We wanted to test if the Neurog3 mutants identified in humans were able to bind to the DNA of target gene promoters, and applied the EMSA setup described section 7.4. This assay tests the ability of Neurog3 to bind to the E-box 1 of the Neurod1 promoter.

Results
Neurog3 generated in the presence of E12 showed a nice band on the EMSA (figure 5.1). Neurog3<sub>R93L</sub> and Neurog3<sub>R107S</sub> also showed a band on the EMSA when generated in the presence of E12 whereas a band for Neurog3<sub>R93L,R107S</sub> generated in the presence of E12 was not detected (figure 5.1).

No bands were detected for Neurog3 or any of the human Neurog3 mutants generated in the absence of E12 or in the presence of wild type Neurog3 (figure 5.1 and data not shown).
Figure 5.1. DNA binding of the human Neurog3 mutants. EMSA showing that Neurog3R93L and Neurog3R107S bind to the DNA probe when generated in vitro in the presence of E12. Neurog3R93L, R107S did not bind to the DNA probe or did so at an undetectable level in this setup.

Conclusions

The human Neurog3 mutants Neurog3R93L and Neurog3R107S generated in the presence of E12 bound to the DNA probe, whereas the double mutant Neurog3R93L, R107S did not bind DNA or did so at an undetectable level. None of the mutants bound DNA at detectable levels in the absence of E12. The bands for Neurog3R93L and Neurog3R107S appeared weaker than the band for Neurog3, and the band for Neurog3R93L seemed weaker than the band for Neurog3R107S. This corresponds well to the findings from luciferase assays where the activity of both mutants was impaired but the activity of Neurog3R93L was lower than the activity of Neurog3R107S (Jensen et al., 2007). Both mutants showed some activity in chicken assays (Jensen et al., 2007). Neurog3R93L, R107S did not show any activity in luciferase assays or chicken assays (Jensen et al., 2007). Both of the mutations were located in the bHLH region of Neurog3, Neurog3R93L in the basic region and Neurog3R107S in the Helix1 region, which are involved in DNA binding and dimerization, respectively (Ma et al., 1994; Sommer et al., 1996). This fact made it plausible that they were DNA-binding mutants, or that the mutations interfered with E-protein interaction, and as showed in another section of this report this interaction is essential for DNA binding (section 3.5).
5.2 *In vivo* activity of the human Neurog3 mutants in the presence of E12

**Purpose**

It was shown that *in vivo* the Neurog3 mutants Neurog3\(^{R93L}\), Neurog3\(^{R107S}\) were active although they showed reduced activity compared to the wild type protein in the luciferase assay, and the double mutant Neurog3\(^{R93L,R107S}\) also had reduced activity in the luciferase assay and was inactive *in vivo* (Jensen et al., 2007). We wanted to investigate if the presence of E12 would restore the activity of Neurog3\(^{R93L,R107S}\) *in vivo*.

**Results**

Chicken embryos were electroporated with E12 or co-electroporated with E12 and Neurog3\(^{R93L,R107S}\) or Neurog3\(^{R107S}\) as described in section 7.10-7.12. As previously described E12 induced some clustering of the electroporated cells by itself (figure 5.2 A and A’). Cells electroporated with Neurog3\(^{R107S}\) and E12 delaminated to some extent supporting earlier findings that Neurog3\(^{R107S}\) was active (figure 5.2 C and C’, (Jensen et al., 2007)). Neurog3\(^{R93L,R107S}\) co-electroporated with E12 induced clustering but not migration of the electroporated cells. These effects resembled the ones observed when embryos were electroporated with E12 alone (figure 5.2 B and B’).

![Figure 5.2](image.png)

**Figure 5.2.** *In vivo* activity of Neurog3\(^{R93L,R107S}\) and Neurog3\(^{R107S}\) in the presence of exogenous E12. Chicken embryos *in ovo* co-electroporated with E12 and Neurog3, Neurog3\(^{R93L,R107S}\) or Neurog3\(^{R107S}\) were subjected to whole mount immunofluorescence staining for glucagon and confocal microscopy. The upper panel pictures represent projections of the confocal sections of a chicken embryo into one picture and the lower panel pictures show optical sections from a representative chicken embryo.
Conclusions

The \textit{in vivo} activity of Neurog3$^{R93L, R107S}$ was not rescued by co-electroporation with E12.

\begin{tcolorbox}
\textbf{Co-authorship publication 1}

\textbf{Mutant neurogenin-3 in congenital malabsorptive diarrhea.}
Authors: Jan N. Jensen, Louise C. Rosenberg, Jacob Hecksher-Sørensen, Palle Serup

It was highly relevant to characterise the human Neurog3 mutants for their activity and function in order to investigate the mechanism by which endocrine cells could develop in the pancreas of the patients affected by the Neurog3 mutants. Neurog3$^{R93L}$ and Neurog3$^{R107S}$ both showed impaired activity on the Neurod1 promoter in luciferase assays, but they had retained the ability to induce glucagon expression and migration of \textit{in ovo}-electroporated cells. Although the lower activity of the human Neurog3 mutants was not sufficient to induce endocrine differentiation in the gut of the patients, it was sufficient to induce endocrine differentiation in the pancreas. Together with the ability of \textbeta{}–cells to proliferate in response to increased demand for insulin this could explain the presence of a \textbeta{}–cell mass that precluded glucose intolerance at least at a young age of the patients carrying these Neurog3 mutants (Kulkarni et al., 2004; Jensen et al., 2007).

The article was published in New England Journal of Medicine as a comment to the original article that described the patients (Wang et al., 2006). The publication is included as appendix 2. My contribution to the publication was described previously in section 5.
\end{tcolorbox}
6. Neurog3 as an activator of Myt1 transcription

At the time this project was in the beginning Myt1 was a newly discovered transcription factor in endocrine development of the pancreas (Gu et al., 2004). In collaboration with Guoqiang Gu from Vanderbilt University we wanted to investigate if Neurog3 regulated Myt1 expression. It was discovered that Neurog3 and Myt1 form a feed-forward expression loop to promote endocrine differentiation in the pancreas (Wang et al., 2008). I contributed to this work with the luciferase assays in figure 6D and E of the publication, which showed that Neurog3 induced transcription from the Neurod1 and Myt1 promoter in 3T3 and P19 cells (the publication is included as appendix 3). As a Neurog3 downstream target Myt1 was interesting at the time, as a downstream transcription factor that would a potentially induce β-cell specification. On that background it was interesting to investigate the function of Myt1 further. It was discovered that Myt1 existed as two splice variants one with a 6-zinc finger motif, and one with a 7-zinc finger motif (Matsushita et al., 2002). The relationship between these two proteins would also be of interest if they had different functions in the developing pancreas. I worked on the Myt1 project for a while until it was established that Myt1 was not essential for development of β-cells and the project was terminated (Wang et al., 2007). My contributions to these investigations are presented in the following section. I investigated the cellular location of the Myt1 splice variants, the dimerization ability of the two proteins, and tested the response on the Myt1 promoter of the Neurog3 truncation mutants described previously in this report. Myt1 expression constructs and the My11 reporter construct were all cloned by Dr. J Hecksher-Sørensen.

6.1 Investigations on the Myt1 protein isoforms “6-zinc finger” and “7-zinc finger”

Purpose
Transcription of the Myt1 gene can give rise to two protein isoforms due to the presence of alternative promoter sites, one containing seven zinc fingers and another containing six zinc fingers. It was interesting to investigate if there was a difference between these isoforms in terms of cellular localisation and if they were able to dimerize. These features may be important for the function of Myt1.

Results
3T3 cells were transfected with HA-tagged Myt1 6-zinc finger, myc-tagged Myt1 7-zinc finger or both constructs. Two days later the cells were fixated and subjected to immunofluorescence staining and microscopy as described in section 7.8. Myt1 6 zinc-finger was predominantly
located in the cytoplasm, whereas Myt1 7-zinc-finger was predominantly located in the nucleus (figure 6.1 A, B, C, D, E, and F). Upon co-transfection the 6- and 7-zinc finger variants retained their subcellular localisation (figure 6.1 G and H).

Figure 6.1. Cellular localisation and dimerization of Myt1 6-zf and Myt1 7-zf. A-H. Cellular localisation of Myt1 6-zinc finger and Myt1 7-zinc-finger in transfected 3T3 cells. Pictures were obtained using a 25x objective. I. Immunoprecipitation of the HA-tag (Myt1 6-zinc finger) from lysates of 3T3 cells transfected with HA-tagged Myt1 6-zinc finger and myc-tagged Myt1 7-zinc finger subjected to western blotting for the myc-tag (Myt1 7-zinc finger). J. Immunoprecipitation of the myc-tag (Myt1 7-zinc finger) from lysates of 3T3 cells transfected with HA-tagged Myt1 6-zinc finger and myc-tagged Myt1 7-zinc finger subjected to western blotting for the HA-tag (Myt1 6-zinc finger).

For immunoprecipitation studies 3T3 cells were transfected with HA-tagged Myt1 6-zinc finger, myc-tagged Myt1 7-zinc finger or both constructs. Two days later the cells were harvested and samples were sonicated, examined for the protein content and subjected to immunoprecipitation as described in section 7.7. The HA-tag and myc-tag were detectable in lysates of co-transfected cells that was not immunoprecipitated but not in untransfected cells (figure 6.1 I and J lanes 1, 3, 5 and 6). Controls showing the presence of the HA-tag in lysates immunoprecipitated for the HA-tag and of the myc-tag in lysates immunoprecipitated for the myc-tag were positive (figure 6.1 I and J, lanes 4). When the lysates were immunoprecipitated for the myc-tag and western blottet for the HA-tag a band appeared on the membrane showing that the Myt1 splice variants were able to interact (figure 6.1 I, lane 2). A band also appeared when lysates were immunoprecipitated for the HA-tag and western blottet for the myc-tag (figure 6.1 J, lane 2) confirming this.
Conclusions
The Myt1 variants locate to different cellular compartments when transfected alone but also when co-transfected. The immunoprecipitation study showed that the splice variants were able to interact, but since immunofluorescence staining of transfected cells showed different localisation, such interaction would be anticipated not to be widespread under natural circumstances.

6.2 In vitro response on the Myt1 promoter of the truncated Neurog3 proteins

Purpose
As Myt1 was identified as an endocrine differentiation transcription factor it would be relevant to place it in the transcription factor hierarchy. More precisely, to understand its function it would be necessary to investigate whether it operated downstream of Neurog3 or if it was activated independently of Neurog3.

Results
Luciferase assays to the Myt1 promoter were optimized as described in section 7.9, and cells were co-transfected with the reporter and the truncated Neurog3 proteins. This would show if the Myt1 promoter would respond to Neurog3, but also if it would respond in the same way as the Neurod1 reporter with regard to the truncated Neurog3 proteins. Neurog3 was able to induce transcription around two-fold from the Myt1 promoter in 3T3 cells and around 5-fold in P19 cells ($p < 0.05$, figure 6.2 A and B). This was also commented in the publication attached as appendix 3 in figure 6D and E. The values in that figure were slightly different from the numbers presented in figure 6.2 A and B because the experiments had been repeated less times for the publication. In 3T3 cells only Neurog3, Neurog3$^{1-178}$, and Neurog3$^{29-214}$ were able to induce a response on the Myt1 promoter (figure 6.2 A). Only the reporter alone, Neurog3$^{1-138}$, Neurog3$^{1-124}$, and Neurog3$^{74-138}$ were significantly different from Neurog3 (figure 6.2 A). The fact that the rest of the truncated Neurog3 proteins were not significantly different from the basal activity level of the reporter or Neurog3 is suggested to be a consequence of the generally low induction values obtained in the 3T3 cell line. In P19 cells, Neurog3, Neurog3$^{1-178}$, Neurog3$^{1-150}$, Neurog3$^{29-214}$, Neurog3$^{56-214}$, Neurog3$^{74-214}$, and Neurog3$^{N89D}$ were able to induce a response from the Myt1 promoter, whereas Neurog3$^{1-138}$, Neurog3$^{1-124}$, Neurog3$^{91-214}$, and Neurog3$^{74-138}$ were not (figure 6.2 B). Neurog3$^{1-178}$, Neurog3$^{1-150}$, Neurog3$^{74-214}$, and Neurog3$^{N89D}$ induced a response that was different to the level induced by Neurog3 (figure 6.2 B). On the Neurod1 promoter only Neurog3$^{1-150}$ was able to induce a response on the promoter.
that was significantly different to the wild type level (figure 6.2 B). It was surprising that Neurog3<sup>N89D</sup> induced a minor response on the Myt1 promoter.

Fig. 6.2. Activity of the Neurog3 truncation mutants on the Myt1 promoter. A. In 3T3 cells the truncated Neurog3 proteins showed quite low induction values and the level of Neurog3 of around 4.5 that was seen in the optimization experiments could not be reproduced. However, numbers were high enough to reflect statistical differences. B. In P19 cells Neurog3 induced the reporter around 5 fold in average. The values represent the response of the activators in relation to the basal activity of the reporter (fold induction), n= 4-8, bars represent standard error of the mean, * marks p < 0.05 compared to the basal activity level of the reporter, dots mark p ≤ 0.05 compared to Neurog3.

Conclusions
With regard to the Myt1 reporter, the P19 cell line again showed higher values than the 3T3 cell line. The Myt1 reporter never reached the high values of the Neurod1 reporter, but a response was clearly induced by Neurog3. The truncated Neurog3 proteins showed approximately the same response pattern as on the Neurod1 reporter except that three in stead of one of the proteins induced lower levels of activity than Neurog3. This suggests that promoter specific differences may apply to the Neurog3 response genes. Perhaps the Neurog3 DNA binding motif or binding partners varies somewhat between the promoters.
6.3 The effect of exogenous E12 on Neurog3 mediated transcription of Myt1

Purpose
As there seems to be some differences between the Neurod1 and the Myt1 promoter response it would be interesting to investigate if this difference also applies to the role of E12. That is, since the addition of exogenous E12 did not increase the Neurog3 response on the Neurod1 promoter would it do so on the Myt1 promoter?

Results
P19 cells were transfected with the Myt1 reporter, Neurog3 (60 or 20 ng per well), and E12 (20 or 60 ng per well), the cells were checked for transfection efficiency two days later and harvested for analysis as described in section 7.9. E12 alone (20 ng) did not have any transcriptional activity on the Myt1 reporter. Neurog3 alone (60 ng) induced a response of approximately 4-fold compared to the basal level of the reporter (figure 6.3). Addition of E12 (20 ng) did not increase the activity of Neurog3. A slight decrease was observed, although this was not significant, which had also been the case for the Neurod1 reporter (figure 6.3). When Neurog3 was co-transfected with a surplus of E12 (20 ng Neurog3 and 60 ng E12) this also tended to decrease the activity of Neurog3 compared to the situation where cells were transfected with equal amounts of Neurog3 and E12 (20 ng of each) (figure 6.3). Again this was not significant.
Figure 6.3. In vitro activity of Neurog3 co-transfected with E12 on the Myt1 promoter.
Luciferase assays were performed with the Myt1 promoter, Neurog3 (20 or 60 ng per well), and E12 (20 or 60 ng per well). The values represent the response of the activators in relation to the basal activity of the reporter (fold induction). Bars represent standard error of the mean, * marks $p < 0.05$ compared to the basal activity level of the reporter, lines indicate no statistical difference between the two values, n=8-9.

Conclusions
Addition of E12 to the luciferase assay did not increase the activity of Neurog3 on the Myt1 promoter. A non-significant tendency of E12 to decrease Neurog3 activity was observed. This was also the case when this setup was tested on the Neurod1 promoter. Again, a suggestion to this is that the endogenous level of E12 was sufficient for Neurog3 dimerization and activity, and that the tendency of exogenous E12 to decrease the Neurog3 response may be a result of sequestering of general transcription factors from the Neurog3/E12 complexes by E12 or E12-complexes with proteins other than Neurog3. See also section 4.3.

Co-authorship publication 2
MyT1 and Ngn3 form a feed-forward expression loop to promote endocrine islet cell differentiation
Authors: Sui Wang, Jacob Hecksher-Sorensen, Yanwen Xu, Aizhen Zhao, Yval Dor, Louise Rosenberg, Palle Serup, Guoqiang Gu.

Guoqiang Gu’s laboratory had generated transgenic mouse lines in which the expression of Myt1 or Neurog3 could be induced in Pdx1 expressing cells at the administration of doxycycline. Investigations with these mouse lines showed that ectopic Myt1 expression in Pdx1 positive cells increased the number of glucagon expressing cells and that this took place through increased expression of Neurog3. Upon ectopic activation of Neurog3 in Pdx1-expressing cells the Myt1 production was dramatically increased. Furthermore, Myt1 expression was reduced to around 5% of the wild type level in Neurog3/-/- mouse, suggesting that Myt1 expression relied largely, but not totally on Neurog3 expression. At the Hagedorn Research Institute Dr. J Hecksher-Sørensen investigated the Myt1 promoter for conserved regions and cloned a 1.8 kb bit of the promoter to be used in front of the firefly luciferase gene as a reporter in luciferase assays. It was determined that Neurog3 was able to induce transcription from the Myt1 promoter. Overall, the investigations suggested that Neurog3 and Myt1 formed a feed-forward loop to enhance the expression of each other.

The article was published in Developmental Biology and is included as appendix 3 (Wang et al., 2008). My contribution to the publication was described previously in this section.
7. Methods

In this section the methods used to obtain the mentioned results are described. Some of the methods are also mentioned in the manuscript included as appendix 1.

7.1 Generation of the Neurog3 mutants

The template DNA used for the clonings was kindly provided by Dr. François Guillemot: Mouse Neurog3 N-terminally tagged with haemaglutinin (HA). The cloning of the Neurog3 mutants was performed using pcr with the primers listed in table 7.1 and 7.2 and was also described in the manuscript enclosed as appendix 1. A myc-tagged Neurog3 was also generated.

The Neurog3 truncation mutants were generated using pcr with the primers listed in table 7.1. The pcr reaction was performed with Phusion, High-Fidelity DNA polymerase (F-530L, New England Biolabs) as described by the manufacturer with the following program on a PTC-200 Peltier Thermal Cycler (MJ Research):

Phusion pcr program:
1. 98°C, 2 minutes (denaturing)
2. 98°C, 10 seconds (denaturing)
3. 52°C-75°C gradient (annealing)
4. 72°C, 10 seconds (extension)
5. goto 2 (29 times)
6. 72°C, 5 minutes (extension)
7. End

The reactions were applied to a 1% agarose gel supplied with ethidium bromide and the most defined band was cut out and gel purified by centrifugation in a spin column at 5000 g for 10 minutes. A little of the resulting DNA was applied to a new 1% agarose gel to check that the DNA was purified and not lost in the process. The Phusion DNA polymerase also has exonuclease activity and produces blunt end DNA fragments.
Table 7.1. Primers used in the cloning of the Neurog3 truncation mutants.

<table>
<thead>
<tr>
<th>Neurog3 construct</th>
<th>Forward primer</th>
<th>Reverse primer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neurog3</td>
<td>5'-CACCATGGGCTACCCATACG-3'</td>
<td>5'-TCACAAGAAGTCTGAGAACAC-3'</td>
</tr>
<tr>
<td>Neurog31-178</td>
<td>5'-CACCATGGGCTACCCATACG-3'</td>
<td>5'-CTACTGGGACGTGGGAGACT-3'</td>
</tr>
<tr>
<td>Neurog31-150</td>
<td>5'-CACCATGGGCTACCCATACG-3'</td>
<td>5'-CTACTGGGACGTGGGAGACT-3'</td>
</tr>
<tr>
<td>Neurog31-138</td>
<td>5'-CACCATGGGCTACCCATACG-3'</td>
<td>5'-CTACTGGGACGTGGGAGACT-3'</td>
</tr>
<tr>
<td>Neurog31-124</td>
<td>5'-CACCATGGGCTACCCATACG-3'</td>
<td>5'-CTACTGGGACGTGGGAGACT-3'</td>
</tr>
<tr>
<td>Neurog329-214</td>
<td>5'-CACCATGGGCTACCCATACG-3'</td>
<td>5'-CTACTGGGACGTGGGAGACT-3'</td>
</tr>
<tr>
<td>Neurog356-214</td>
<td>5'-CACCATGGGCTACCCATACG-3'</td>
<td>5'-CTACTGGGACGTGGGAGACT-3'</td>
</tr>
<tr>
<td>Neurog374-214</td>
<td>5'-CACCATGGGCTACCCATACG-3'</td>
<td>5'-CTACTGGGACGTGGGAGACT-3'</td>
</tr>
<tr>
<td>Neurog391-214</td>
<td>5'-CACCATGGGCTACCCATACG-3'</td>
<td>5'-CTACTGGGACGTGGGAGACT-3'</td>
</tr>
<tr>
<td>Neurog374-138</td>
<td>5'-CACCATGGGCTACCCATACG-3'</td>
<td>5'-CTACTGGGACGTGGGAGACT-3'</td>
</tr>
</tbody>
</table>

Introduction of the point mutation N89D into Neurog3 was first attempted using the QuickChange Site-Directed Mutagenesis Kit (200518, Stratagene) without any luck. In stead it was introduced into Neurog3 using chimaeric PCR. In the first step two PCR fragments were generated both consisting of a piece of truncated Neurog3 with the point mutation (table 7.2). The two pieces were mixed and reannealed using the following program on a PTC-200 Peltier Thermal Cycler (MJ Research):

Chimaeric PCR program:
1. 98°C, 2 minutes (denaturing)
2. 98°C, 10 seconds (denaturing)
3. 98°C, 10 seconds (annealing)
4. 98°C, 10 seconds (extension)
5. goto 2, 9 cycles
6. 4°C, 30 minutes
7. End

Afterwards the chimaeric PCR products were amplified using the primers for the full length Neurog3 with the Phusion PCR program, the product being a full length Neurog3 with the point mutation. The products of the PCR reactions were run on a 1% agarose gel supplied with
ethidium bromide and the bands for the full length Neurog3 with point mutations were cut out and gel purified as described for the truncation mutants. The primers are listed in table 7.2.

<table>
<thead>
<tr>
<th>Neurog3 construct</th>
<th>Forward primer</th>
<th>Reverse primer</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Step 1</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neurog3^{1-94,N89D}</td>
<td>5’-CACCATGGGCTACCCATACG-3’</td>
<td>5’-CGGCCGAAGAAGGCCGATGATCGCGAGCGCAAT-3’</td>
</tr>
<tr>
<td>Neurog3^{84-214,N89D}</td>
<td>5’-ATTGCCTCGCGATCATCGGCCTTCTTGCCGCG-3’</td>
<td>5’-TCACAAGAAGTCTGAGAACAC-3’</td>
</tr>
<tr>
<td><strong>Step 2</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neurog3^{N89D}</td>
<td>5’-CACCATGGGCTACCCATACG-3’</td>
<td>5’-TCACAAGAAGTCTGAGAACAC-3’</td>
</tr>
</tbody>
</table>

**Table 7.2.** Primers used for introducing the point mutation into Neurog3.

### 7.2 Gateway cloning

The Gateway system offers a quick and easy way to clone the gene of interest directionally into your vector as long as it contains a Gateway cassette. First the PCR product is cloned into an entry clone in which it will insert in the right direction in most instances because of the pairing of a CACC sequence of the PCR product and a GTGG sequence of the vector. The entry vector also contains attL sites that will recombine with attR sites of a destination vector. Cloning of the PCR product into the entry vector enables you to clone the gene of interest further on into different destination vectors quick and easily.

The PCR fragments of the Neurog3 mutants were cloned into the pENTR™/D-Topo® vector that accepts blunt end PCR fragments and transformed into One Shot® TOP10 chemically competent cells (K2400-20, Invitrogen). The bacteria were plated on kanamycin containing agar and five-eight colonies per construct were picked and checked for correct insertion of the PCR-fragment using the CloneChecker™ System (11666-013, Invitrogen). Two clones per construct that contained the PCR insert were picked, propagated in selective LB-media, and plasmids were purified using the QIAprep Spin Miniprep Kit (27106, QIAGEN) and sent off for sequencing at Eurofins MWG Operon (Anzingerstr. 7a, 85560 Ebersberg, Germany, www.eurofins.com).

DNA from a clone with the correct gene sequence was cloned into the destination vector pDEST-CAGGS-IRES-EGFP (pCGIG) using Gateway® LR Clonase™ Enzyme Mix (11791-043, Invitrogen) and transformed into One Shot® TOP10 chemically competent cells (K2400-20,
Invitrogen). The pCGIG destination plasmid was kindly provided by Dr. Anne Grapin-Botton, who had modified it to contain the Gateway cassette (figure 7.1). Bacteria were propagated in ampicillin containing LB-media (1:1000) and plasmids were purified using the QIAGEN Plasmid Mega Kit (12183, Qiagen). The purified plasmids were used for transfection of cells or in ovo electroporation. All constructs were also cloned from the pENTR™/D-TOPO® vector into the pEXP2-DEST vector (V960-02, Invitrogen). This destination vector encodes a T7 promoter that can be used with the TNT®T7 Quick Coupled Transcription/Translation System (L1171, Promega) to generate proteins in vitro. The cloning procedure was the same as for the pCGIG destination vector with the exception that the resulting plasmids were transformed into One Shot® ccdB Survival™ 2 T1 Chemically Competent Cells (A10460, Invitrogen). This bacteria strain is resistant to the cell division inhibitory effects of the ccdB gene product encoded by the pEXP2-DEST vector.

---

Figure 7.1. The pCGIG destination vector. The pCGIG plasmid was used as the destination vector for all Neurog3 constructs. The name reflects the contents as the letters are related to p (plasmid), C (Caggs), G (Gateway), I (Ires), G (GFP). The gene of interest was inserted between attR1 and attR2 using Gateway cloning as described in the text. The plasmid map was modified from the computer software Vector NTI (Invitrogen).

7.3 In vitro generation of proteins

Binding of proteins to DNA is often analysed using electrophoresis mobility shift assay (EMSA). This assay allows you to use in vitro generated proteins and save the labour and time otherwise used to prepare DNA, transf ect cells, and make proper nuclear lysates. The TNT®T7
Quick Coupled Transcription/Translation System (L1171, Promega) was used to generate proteins \textit{in vitro}. This system contains all components necessary to mediate transcription and translation of the gene of interest when added in a plasmid containing the T7 promoter. The Neurog3 constructs were cloned into the pEXP2-DEST vector for this purpose as described in section 7.2. Plasmids were used in the concentration 0.5 µg/µl and for one reaction 2 µl plasmid was added. When two proteins were generated in one reaction 1 µl of each plasmid was added. The reaction incubated in 60-90 minutes at 30 °C and protein was ready to use. It was not possible to measure the protein content on the Biophotometer because of the red colour of the TNT® Quick Master Mix.

\textbf{7.4 EMSA}

The EMSAs were performed in order to investigate the ability of the Neurog3 mutants to bind to DNA. The protein was generated \textit{in vitro} as described above and the DNA probe was designed from the promoter of the Neurog3 downstream target gene Neurod1.

Designing and preparing the probe
It was previously published that Neurog3 binds to the CANNTG sequence of target gene promoters and that in the Neurod1 promoter Neurog3 binds to the CANNTG-containing E-box 1 very well (Huang et al., 2000). The sequence of this E-box was used as a template for the probe. The CANNTG sequence was flanked by two bases in each end and BamHI and BglII restriction sites giving a probe of 20 bases. The sequence of the forward oligo was 5’-GATCACCATATGGCA-3’ and of the reverse oligo 5’-GATCTGCCATATGGTG-3’. This would leave an overhang of 4 bases in each end. The oligos were ordered from at Eurofins MWG Operon (Anzingerstr. 7a, 85560 Ebersberg, Germany, www.eurofins.com). The forward and reverse oligos were diluted to 100 pmol/µl, mixed in equal amounts (20 µl of each), and after boiling in 60 µl water they were left at room temperature to cool down and anneal. The probe was prepared by adding $\alpha$-$^{32}$P-dATP (AA0074, Amersham) and nucleotides to the annealed oligos using the Klenow polymerase (M2201, Promega) to label the probe.

Optimising the EMSA
In order to achieve a clear EMSA, several approaches were tried. We tried different ways to purify the probe, different EMSA buffers to optimise protein-DNA binding (table 7.3), and adjusted the amount and composition of Poly(dIdC)-Poly(dIdC) and Poly(dAdT)-Poly(dAdT) to reduce unspecific binding of proteins to the probe.
To purify the probe we tested purification over a Nick column (17-0855-02, Pharmacia), purification over a spin column (30 minutes at 14000g, Ultracel YM-10, Millipore), a combination of gel purification and the spin column, and obtaining the Nick column yield in fractions to identify the fraction with largest amount of probe. For gel purification the probe was run on a 4-20% TBE gel (EC6225, Invitrogen) in 1x TBE buffer (LC6675, Invitrogen). The band was cut out and the bit of gel was mashed and left to shake in a gel extraction buffer (500 mM NaAc, 2 mM EDTA) to extract the DNA. The mash was centrifuged at the maximum g for 10 minutes to remove gel material from the supernatant containing the probe. None of the tested purification methods improved the appearance of the EMSAs and purification over a Nick column was selected for the assays, as this method was routinely used in our laboratory.

To reduce unspecific binding of the probe we tested different amounts of the nonspecific competitor Poly(dIdC)-Poly(dIdC) or replaced it with a mix of Poly(dIdC)-Poly(dIdC) and Poly(dAdT)-Poly(dAdT). This had no effect on the appearance of the bands on the gel (data not shown), and 10% of Poly(dIdC)-Poly(dIdC) was used as this was routinely used in our laboratory.

Three EMSA buffers were tested to optimise probe-protein binding (table 7.3). EMSA buffer 1 was recommended by the people in house doing EMSA on a regular basis. EMSA buffer 2 was recommended by a colleague in the department who previously optimised it for bHLH factor binding to DNA. EMSA buffer 3 was recommended by a colleague with extensive EMSA experience on the basis of the results with the other two buffers. The bands on the EMSA appeared the same regardless of the EMSA buffer used, and EMSA buffer 1 was used for the assays as this was routinely used in our laboratory.

<table>
<thead>
<tr>
<th>Contents</th>
<th>EMSA buffer 1 (Sif)</th>
<th>EMSA buffer 2 (HVP)</th>
<th>EMSA buffer 3 (FRGA)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HEPES pH 7.9 (Sigma)</td>
<td>100 mM</td>
<td>125 mM</td>
<td>100 mM</td>
</tr>
<tr>
<td>NaCl (Merck)</td>
<td>50 mM</td>
<td>750 mM</td>
<td>50 mM</td>
</tr>
<tr>
<td>KCl (Merck)</td>
<td>5 mM</td>
<td>5 mM</td>
<td>5 mM</td>
</tr>
<tr>
<td>MgCl₂ (Merck)</td>
<td>50%</td>
<td>50%</td>
<td>50%</td>
</tr>
<tr>
<td>EDTA (Merck)</td>
<td>50%</td>
<td>25 mM</td>
<td>50%</td>
</tr>
<tr>
<td>Glycerol (Merck)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DTT (GibcoBRL)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 7.3. EMSA buffers tested.
If the EMSA should have been optimised further, I would have tried purifying the *in vitro* generated Neurog3-proteins before adding them to the binding reaction to reduce unspecific binding to the probe of other components in the TNT® Quick Master Mix. This was however not tested.

On the basis of the experiments testing different EMSA buffers and varying the amount and composition of the non-specific competitor the following procedure was applied for the assays:

The samples were prepared by mixing the following components:

- 6.4 µl *in vitro* generated protein mix
- 2.4 µl 5xEMSA buffer (EMSA buffer 1)
- 1.2 µl Poly(dIdC)-Poly(dIdC) (27-7880-01, Amersham)
- 2.0 µl probe (20 fmol)

In competition assays 200x unlabelled annealed oligo was added to the sample. In supershift assays antibodies were added and the samples incubated for 30 minutes at 4°C before the probe was added. Samples incubated for 30 minutes at 30°C for the binding reaction to occur, dNTPs (500 µM) were added to reduce unspecific binding, and EMSA loading buffer (LC6678, Invitrogen) was added to the samples before loading. The samples were loaded on 6% DNA retardation gels (EC63655, Invitrogen) and run in 0.5x TBE buffer (LC6675) at 100 V for 1 hour. The gel was transferred to whatman filter paper (3 mm chromatography paper, 3030917, Whatman Int., Ltd.), covered with stretch cling film (food wrapping), and dried for about an hour at 80°C in a vacuum-dryer (Model 583, Biorad). The EMSA was developed for 1-9 days and visualised using a Typhoon 8600 scanner. The specificity of the Neurog3-E12 band was determined by supershift with 2 µg antibody against the HA-tag (HA.11, MMS-101P, Nordic Biosite) or 2 µg antibody against E12 (39-0200, Zymed Laboratories) and competition with unlabelled probe as described (figure 7.2). The myc-tagged Neurog3 was also tested for DNA binding in the presence of E12 and the specificity tested with E12 antibody or 200x oligo.
Neurog3 did not bind to DNA at detectable levels in the absence of E12 in this setup. Neurog3 generated in the presence of E12 gives a nice band that can be supershifted with antibody against the HA-tag, however, the antibody against E12 didn’t seem to work. Adding 200x unmarked oligo reduced the strength of the band but also removed the unspecific band below. Neurog3\textsuperscript{N89D} did not bind to the probe or did so at an undetectable level in this setup, neither in the absence or presence of E12. Myc-tagged Neurog3 bound to the probe in the presence but not in the absence of E12 and adding 200x unmarked oligo reduced both the specific and the unspecific band. Again the E12 antibody doesn’t seem to work on EMSA.

7.5 Cell cultivation

3T3 mouse embryonic fibroblast cells were transfected for immunostainings and for preparing lysates for SDS-PAGE and western blotting and for immunoprecipitations. P19 mouse teratocarcinoma cells were transfected for luciferase assays. The 3T3 cells were grown in DMEM (31885-023, Invitrogen) supplemented with 10 % heat inactivated foetal bovine serum (26140079, Invitrogen), 100 U/ml penicillin (15140122, Invitrogen), and 100 µg/ml streptomycin (15140122, Invitrogen), and propagated once a week at 1:200. The P19 cells were grown in MEM Alpha Medium (32571-028, Invitrogen) supplemented with 10 % heat inactivated foetal bovine serum (26140079, Invitrogen), 100 U/ml penicillin (15140122, Invitrogen), and 100 µg/ml streptomycin (15140122, Invitrogen), and propagated twice a week at 1:100 or 1:200. Cells were transfected using Lipofectamine2000 (11668-019, Invitrogen) as described by the manufacturer. For immunofluorescence stainings 3T3 cells
were seeded in 8 chamber slides at a density of 12,500 cells per well in 200 µl media. For immunoprecipitation lysates 3T3 cells were seeded in T175 flasks. For SDS-PAGE lysates 3T3 cells were seeded in T80 flasks. Cells were transfected at a density of approximately 75% and harvested for lysates or fixated for immunofluorescence stainings two days later. For luciferase assays P19 cells were seeded in 96 well plates at a density of 20,000 cells per well in 100 µl media and 3T3 cells were seeded at a density of 8000 cells per well. Transfection efficiency was approximately 30-50% in P19 cells and 10-30% in 3T3 cells as judged by the fraction of GFP expressing cells two days after transfection. The details on seeding of cells for transfections are listed in table 7.4.

<table>
<thead>
<tr>
<th>Cell line</th>
<th>Culture flask</th>
<th>Cells seeded</th>
<th>Media</th>
<th>DNA used for transfection</th>
</tr>
</thead>
<tbody>
<tr>
<td>3T3</td>
<td>T80</td>
<td>-</td>
<td>15 ml</td>
<td>21000 ng</td>
</tr>
<tr>
<td></td>
<td>T175</td>
<td>-</td>
<td>30 ml</td>
<td>46000 ng</td>
</tr>
<tr>
<td></td>
<td>8-chamber slide</td>
<td>12,500 cells/well</td>
<td>200 µ/well</td>
<td>250 ng</td>
</tr>
<tr>
<td></td>
<td>96-well plate</td>
<td>8000 cells/well</td>
<td>100 µ/well</td>
<td>100 ng</td>
</tr>
<tr>
<td>P19</td>
<td>T80</td>
<td>-</td>
<td>15 ml</td>
<td>21000 ng</td>
</tr>
<tr>
<td></td>
<td>T175</td>
<td>-</td>
<td>30 ml</td>
<td>46000 ng</td>
</tr>
<tr>
<td></td>
<td>8-chamber slide</td>
<td>25,000 cells/well</td>
<td>200 µ/well</td>
<td>250 ng</td>
</tr>
<tr>
<td></td>
<td>96-well plate</td>
<td>20,000 cells/well</td>
<td>100 µ/well</td>
<td>100 ng</td>
</tr>
</tbody>
</table>

Table 7.4. Transfection of 3T3 and P19 cells. T175 flasks were used for seeding and transfection of cells for immunoprecipitation. T80 flasks were used for seeding and transfection of cells for western blots. 8-chamber slides were used for seeding and transfection of cells for immunofluorescence stainings. 96-well plates were used for seeding and transfection of cells for luciferase assays.

**7.6 SDS-PAGE and Western blotting**

SDS-PAGE and Western blot analysis was performed in order to examine if the modified Neurog3 proteins were expressed in cells. 3T3 cells grown in T80 culture flasks were transfected with the Neurog3 expression plasmids at approximately 75% confluence. Two days later the percentage of EGFP expressing cells was approximated as a measure of transfection efficiency, and the cells were washed with 4°C PBS before they were lysed on ice with 1 ml NP-40 lysis buffer supplemented with 1 mM of the proteinase inhibitor AEBSF (A8456, Sigma-Aldrich) and scraped off the flasks. Lysates were homogenised with a 21 gauge needle, sonicated, and the protein content measured using the Bio-Rad Protein Assay (500-0006, Bio-Rad) and a Eppendorph Biophotometer. The lysates were normalized with regard to protein content and approximately 10-15 µg protein was loaded per well of 10- to 15-wells 10% Bis-Tris gels (NP0303BOX, Invitrogen). SDS-PAGE and Western blotting was also
performed of the truncated Neurog3 proteins generated in vitro. In order to optimize detection of the in vitro generated Neurog3 variants different amounts of protein was loaded on the gel. The samples were supplied with 1/20 DTT and NuPAGE® LDS sample buffer (NP0007, Invitrogen) and boiled for 5 minutes at 90°C to denature the protein before loading. Gels were run in MES buffer (NP0002, Invitrogen) at 100 W, 200 V, 100 mAmp per gel for approximately 45 minutes using the Xcell SureLock™ Mini Cell electrophoresis system (EI0002, Invitrogen). Using the same system the proteins were transferred to nitrocellulose membranes (LC2000, Invitrogen) in NuPAGE® transfer buffer (NP0006, Invitrogen) applying 30V, 170 mAmp for 70 minutes. The transfer was briefly evaluated with ponceau staining and the membranes were blocked for at least 30 minutes in 2% skimmed milk (70166, Fluka, Biochemika) in TBST in order to prevent contamination of the membranes with other proteins. Primary antibodies were diluted in TBST/2% skimmed milk and the membranes incubated for at least one hour at a rocking table. Membranes were washed with TBST 3 times 10 minutes before incubation with secondary antibodies that were diluted 1:10.000 in TBST/2% skimmed milk. Neurog3 was detected with antibodies against the HA-tag (HA.11, 1:500, MMS-101P, Nordic BioSite) and on the cell lysate blots EGFP was detected as a control of vector expression (GFP, 1:1000, M048-3, Nordic BioSite). Secondary antibodies were the HRP-conjugated goat anti-mouse IgG-HRP (1:10.000, sc-2005, Santa Cruz) and the goat anti-rabbit IgG-HRP (1:10.000, sc-2004, Santa Cruz). Bands were visualized using SuperSignal® West Dura Extended Duratin Substrate (37071, Pierce) and a UVP Biospectrum AC Imager system.

7.7 Immunoprecipitation

3T3 cells were seeded in T175 flasks transfected with the relevant plasmids using Lipofectamine 2000 as described in a previous section (section 7.5). For the Myt1 immunoprecipitation one flask was transfected with the “empty vector”, that is the pCGIG vector encoding EGFP, one with the Myt1 6-zinc finger plasmid and “empty vector”, one with the MuT1 7-zinc finger plasmid and “empty vector”, and one with the Myt1 6-zinc finger and the Myt1 7-zinc finger plasmids. Two days later the cells were washed with PBS and harvested using a NP40 lysis buffer supplemented with AEBSF (0.5 M). The samples were homogenised with a 21 gauge needle, sonicated, and centrifuged briefly to remove particulate matter. The samples were stored at -20°C. Antibody was added to 500 µl sample in the dilution suggested by the manufacturers (HA.11, MMS-101P, Nordic Biosite: 1:150, 9e10, SC-40: 2 µg per 500 µl) and the binding reaction incubated for at least one hour at 4°C with gentle rocking. The immunoprecipitation was performed using Protein G Sepharose 4 Fast Flow (17-0618-01, Amersham), which is suited for antibodies raised in mouse. The media was washed three times
with lysis buffer before the slurry of 50% media and 50% lysis buffer was prepared. 50 µl of the slurry was added to the antibody-supplied samples and the reaction incubated for one hour at 4 °C with gentle rocking to allow for the Protein G sepharose – antibody binding reaction. The samples were centrifuged briefly at 12000 g and the supernatant was removed. The beads were washed three times with lysis buffer and once with the wash buffer suggested by the manufacturer. The final pellet was suspended in sample buffer, boiled for 3 minutes to release antibody and antigen from the beads, and centrifuged briefly at 12000 g to remove the beads. Bromphenol blue was added to make samples ready for SDS-PAGE.

7.8 Immunofluorescence staining of cells

The cells were fixated in 4% PFA two days after transfection. The Neurog3 proteins were detected with an antibody against the HA-tag (HA.11, 1:800, MMS-101P, Nordic BioSite) and the secondary antibody was a cy-3 conjugated donkey-anti-mouse antibody (1:200, 715-165-151, Jackson ImmunoResearch). Pictures were obtained on a Zeiss imager Z1 LSM 510 Meta confocal microscope using the Zeiss 25X/0.80 oil objective.

7.9 Luciferase assays

Optimizing the assay

In order to test the in vitro activity of the Neurog3 mutants luciferase assays were set up. In order to test the activity of the Neurog3 mutants on the Neurod1 and the Myt1 promoters the assay first needed to be optimised for a good response of Neurog3 that would allow us to discriminate between minor differences in activity. The Neurod1 reporter plasmid was cloned by Dr. Jacob Hecksher-Sørensen based on a previously described construct and consisted of the 2.2 kb promoter in front of the firefly luciferase gene in the pGL3 basic vector (E1751, Promega) (Huang et al., 2000). The Myt1-reporter plasmid was cloned by Dr. J. Hecksher-Sørensen and consisted of the 1.8 kb Myt1 promoter in front of the firefly luciferase gene in the pGL3 basic vector (E1751, Promega). Firefly luciferase expression would be activated upon Neurog3 binding to the Neurod1 or Myt1 promoters and mediate transcription. The activity of Neurog3 would be determined by the firefly luciferase activity measured on an Orion microplate luminometer (Berthold Detection System). Cells were co-transfected with a renilla luciferase expression plamid (pRL-CMV, E2261, Promega). By normalizing firefly luciferase values to the renilla luciferase values variability in transfection efficiency could be accounted for. Initially the setup was optimized in 3T3 cells testing different amounts of reporter and
activator. First was tested 20 ng reporter with 40 ng activator, 40 ng reporter with 30 ng activator, and 60 ng reporter with 20 ng activator. On the Neurod1 reporter this resulted in a 2-3 fold induction compared to the basal activity of the reporter at all conditions (figure 7.3 A). On the Myt1 promoter we saw a three fold induction of the reporter when there was a surplus of activator to reporter (40 ng activator to 20 ng reporter) whereas the other conditions showed lover to no induction (figure 7.4 A). On that background we lowered the amount of reporter and activator, in order to see if we could get the same response with less plasmid. We tested quadruplicates of 10, 20, and 40 ng reporter with 5, 10, 20, 40, 60, and 80 ng activator (figure 7.3 B and 7.4 B) and decided to continue this assay with 20 ng reporter to 60 ng activator. This combination showed the nicest response values of approximately 4 fold activation on the Neurod1 reporter and 5 fold activation on the Myt1 reporter (figure 7.3 B and 7.4 B). However, after determining the amounts of the plasmids to be used, we also tested the assay in P19 cells, as it would be desirable to get an even higher level of activation if we should be able to determine potential minor differences in activity of the Neurog3 mutants. The P19 cell line showed much nicer induction levels as Neurog3 was able to activate the Neurod1 reporter 40 fold and the Myt1 reporter 5 fold compared to the basal activity of the reporters (figure 3.8 and 6.2). Furthermore, transfection efficiency was as high as 30-50% in P19 cells whereas in the 3T3 cells it was around 10-30% as estimated by visual examination of the fraction of cells expressing GFP using a fluorescence microscope (Olympus, 1x51).

Experiments were continued with both the P19 cell line and the 3T3 cell line.

![Figure 7.3. Optimizing luciferase assays on the Neurod1 reporter. A. 20, 40, or 60 ng of the reporter containing the Neurod1 promoter was tested with 40, 30, and 20 ng of the Neurog3 expression plasmid. All combinations resulted in a three-fold increase in the relative luciferase activity compared to the reporter co-transfected with the empty vector. B. 10, 20, or 40 ng of the reporter containing the]
Neurod1 promoter was tested with 0, 5, 10, 20, 40, 60, and 80 ng of the Neurog3 expression plasmid. n=1.

**Figure 7.4. Optimizing luciferase assays on the Myt1 reporter.**

A. 20, 40, or 60 ng of the reporter containing the Myt1 promoter was tested with 40, 30, and 20 ng of the Neurog3 expression plasmid. The only combination that resulted in a marked increase in the relative luciferase activity compared to the reporter co-transfected with the empty vector (approximately three-fold) was the condition with 20 ng reporter and 40 ng activator.

B. 10, 20, or 40 ng of the reporter containing the Myt1 promoter was tested with 0, 5, 10, 20, 40, 60, and 80 ng of the Neurog3 expression plasmid. n=1.

The procedure of the luciferase assays were as follows: Cells were seeded in 96 well plates and transfected the following day (table 7.4). 100 ng DNA was mixed with the reduced serum media Opti-Mem® (S1985) to a total of 5 µl per well. In parallel, 0,5 µl Lipofectamin2000 (11668-019, Invitrogen) was added to 4,5 µl optimem. The 5 µl DNA/optimem and Lipofectamin2000/optimem mixes were mixed and incubated for 30 minutes at room temperature. The cells were washed with optimem. 40 µl optimem was added to the 10 µl transfection mix and the total of 50 µl transfection media was added to each well of cells. Cells were incubated with the transfection media for 4-6 hours, the media was removed and cell culture media was added. Two days later, the percentage of cells expressing EGFP was approximated as a measure of transfection efficiency and the cells were harvested. The cells were washed with PBS and 100 µl 1x Passive Lysis Buffer (1941, Promega) was added per well. Cells were left on a rocking table for at least 15 minutes before the samples were analysed at an Orion Microplate Luminometer (Berthold Detection Systems) using the Dual-
Luciferase® Reporter Assay System (1960, Promega). The procedure for the luciferase assays was also mentioned in the manuscript enclosed as appendix 1.

7.10 **In ovo** electroporation

*In ovo* electroporation was performed as previously described (Grapin-Botton et al., 2001). Briefly, 5 ml of egg white was removed from the eggs at HH11-15, shells were windowed and DNA was injected under the embryo with a mouth pipette (Hamburger V, 1951). Electrodes were placed longitudinally with the embryos axis with the negative electrode underneath the embryo and the positive electrode above. PBS was added to connect the electrodes on each side of the embryo and 4 pulses of 12 V were applied. The electricity enables the cells to take up the DNA that migrates towards the positive electrode. Embryos incubated for two days at 37°C before harvesting. They were decapitated and placed in Lilly’s fix (4% PFA) over night at 4°C. Embryos were transferred to methanol and could be stored at -20°C until further treatment. The procedure was also described in the manuscript enclosed at appendix 1.

7.11 Immunofluorescence staining of whole mounted chicken embryos

Immunofluorescence staining of the whole mounted chicken embryos was performed as previously described (Ahnfelt-Ronne et al., 2007b). Briefly, the embryos were permeabilised in a methanol gradient each step lasting one hour: 100% methanol, methanol:DMSO (4:1), methanol/DMSO/H₂O₂ (7:2:1), before 5 minutes of rehydration in PBS and one hour of blocking in TNB-buffer. Primary antibodies were applied in 0.5% TNB in PBS at 4 °C overnight at gentle rocking (see table 7.5 for dilutions). The tissue was washed in PBS three times of 10 minutes. Secondary antibodies were applied in 0.5% TNB in PBS for 4 hours at room temperature or over night at 4°C at gentle rocking (see table 7.5 for dilutions). The tissue was washed in PBS three times of 10 minutes before transfer to absolute methanol. The specimens were stored at -20°C.
<table>
<thead>
<tr>
<th>Target</th>
<th>Identity</th>
<th>Origin</th>
<th>Working dilution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary</td>
<td>Antibodies</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GFP</td>
<td>632460, Clontech</td>
<td>Rabbit</td>
<td>1:2000</td>
</tr>
<tr>
<td>Foxa2</td>
<td>SC6554, Santa Cruz</td>
<td>Goat</td>
<td>1:500</td>
</tr>
<tr>
<td>Glucagon</td>
<td>Glo-001, Novo Nordisk, Denmark</td>
<td>Mouse</td>
<td>1:100</td>
</tr>
<tr>
<td>Insulin</td>
<td>Ab7842, Abcam</td>
<td>Guinea pig</td>
<td>1:500</td>
</tr>
<tr>
<td>Secondary</td>
<td>Antibodies</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cy-2-anti rabbit</td>
<td>711-225-152, Jackson ImmunoResearch</td>
<td>Donkey</td>
<td>1:300</td>
</tr>
<tr>
<td>Cy-3-anti-goat</td>
<td>705-165-147, Jackson ImmunoResearch</td>
<td>Donkey</td>
<td>1:300</td>
</tr>
<tr>
<td>Cy-5-anti-mouse</td>
<td>715-175-151, Jackson ImmunoResearch</td>
<td>Donkey</td>
<td>1:300</td>
</tr>
<tr>
<td>Cy-5-anti- guinean pig</td>
<td>706-175-148, Jackson ImmunoResearch</td>
<td>Donkey</td>
<td>1:300</td>
</tr>
</tbody>
</table>

Table 7.5. Antibodies used in immunofluorescence staining of whole mounted chicken embryos.

7.12 Confocal microscopy of immunofluorescence stained chicken embryos

The specimens were cleared in BABB (Benzyl alchohol:Benzyl benzoate, 1:2), mounted in 3 mm depression slides, and scanned at a Zeiss Imager Z1 LSM 510 Meta confocal microscope using a Plan-Neofluar 10x/0.30 objective or a Achroplan 20x/0.50 w PH2 water objective and the LSM software.

7.13 Quantification of differentiation

Differentiation was quantified by counting of red cells expressing glucagon and green cells expressing EGFP. Differentiation was expressed as the percentage of the electroporated cells co-expressing glucagon. Cells were counted posterior to the dorsal pancreas bud and cells that were closer to the bud than a buds-length were not included (figure 7.5). Cells were counted on every second-third optical section corresponding to a distance of 20-25 µm to avoid counting the same cells more than once. The values varied quite a bit between embryos electroporated with the same construct partly because of differences in electroporation efficiency and the quality of the stainings. This was reflected by the quite large standard errors of the mean (figure 3.11).
Figure 7.5. Example of how differentiation was quantified. A. Example of a whole mount immunofluorescence stained chicken embryo electroporated with Neurog3 at HH11-15 and harvested two days later. The picture shows a projection of all optical section into one picture. B. An optical section of the embryo. C. The colour channels were split when cells on the optical sections were counted. This picture shows the red channel, that is the glucagon staining. The white line separates the cells that were counted (posterior) and cells that were not counted (anterior). D. EGFP staining of the same optical section with the line separating cells that were counted and cells that were not counted. Pictures were obtained with the 10x objective.

7.14 Statistics

For both the lucifease assays and the quantification of differentiation was applied a standard students t-test. In both instances the samples were unpaired with a two-tailed distribution and unequal variances. $P<0.05$ was regarded as statistically significant.
8. Discussion

The overall objective of this thesis was to understand in greater detail the role of the transcription factor Neurog3 in the differentiation of pancreatic progenitor cells towards an endocrine fate. It is well established that Neurog3 induces endocrine differentiation, but exactly how this happens is not very well characterised. With this report we tried to address this question from different angles. We investigated if certain Neurog3 protein domains were important for its function in vivo, if its function in vitro and in vivo would be increased by the E-protein E12, the DNA-binding of the human Neurog3 mutants and the effect of E12 on their in vivo activity, and if Myt1 was a direct target of Neurog3. The findings from these investigations are described in this section.

8.1 Neurog3 protein domains

The designed Neurog3 mutants were characterised in order to identify regions that could be essential for Neurog3 function. When examined for their expression in cells using western blotting, the proteins with N-terminal truncations showed weaker bands than the wild type protein and the proteins with C-terminal truncations. This would suggest that they were expressed at a lower level or were degraded more rapidly in cells than the proteins with C-terminal truncations. That may be true to some extent, since these proteins also stained weaker in immunofluorescence stained cells. However, based on the immunofluorescence stainings there is no doubt that the proteins were expressed in cells. This is supported by the fact that they were capable of inducing a transcriptional response in luciferase assays that was comparable to the activity of the wild-type protein. When the proteins were generated in vitro they were not subjected to cellular degradation but the proteins with N-terminal truncations still showed weak bands. This suggests that the antibody detection of the N-terminally located HA-epitope on western blots and in immunofluorescent stained cells may be impaired for these proteins. The sequence just C-terminal to the HA-epitope is different in each of the proteins with N-terminal truncations and they are different from the sequence in the wild type and C-terminally truncated proteins. Perhaps these sequences influence the antibody detection negatively. In order to avoid this, a short linker region could have been introduced, so that the sequence just C-terminally to the tag had been the same for all the proteins.

The nuclear localisation of the truncated proteins was not affected for most of the proteins. However, the proteins with a truncated bHLH domain (Neurog3^{1-124} and Neurog3^{91-214}) were clearly more localised to the cytoplasm than the proteins with an intact bHLH domain. The
protein that lacked the entire C-terminal domain (Neurog31-138) was primarily located in the nucleus, whereas the protein that lacked the entire N-terminal domain (Neurog374-214) was also present to some extent in the cytoplasm although not quite to the same degree as the proteins with truncated bHLH domain. The cellular location of the bHLH domain which lacked both the C- and N-terminal domains resembled that of the protein which lacked the N-terminal domain (Neurog374-214), suggesting that the region just N-terminal to the bHLH domain may be involved in nuclear localisation. This seems plausible, since the Neurod1 basic domain is defined as to include more amino acids N-terminally to the sequence included in the Neurog3 bHLH domain, and were found to contain basic residues involved in nuclear localisation (figure 8.1) (Mehmood et al., 2009). The impaired nuclear localisation of the proteins lacking the N-terminal domain was supported by co-transfecteeion studies with wild type Neurog3 or E12.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure8.1.png}
\caption{The basic-helix-loop-helix domain of Neurog3 and Neurod1. The bHLH domain of Neurod1 is defined to include more amino acids than the Neurog3 bHLH domain. The basic domain is marked in red (Sommer et al., 1996; Mehmood et al., 2009).}
\end{figure}

The EMSA studies showed that the presence of E12 was required for Neurog3 DNA binding and that in the presence of E12, the DNA-binding of truncated protein lacking the N-terminal domain (Neurog374-214) was much stronger than for any of the other proteins including wild type Neurog3. That is remarkable when you take into account the relatively weak bands this protein showed on western blots and that its presence in cells was not restricted to the nuclear compartment. The EMSA confirmed that the DNA binding of Neurog3N89D was impaired, and this was supported by the lack of transcriptional activity in spite of a nuclear localisation. Neurog3N89D was used as a negative control in the experiments described in this report, in the manuscript included as appendix 1, and in two published papers (Jensen et al., 2007; Wang et al., 2008). DNA binding was also absent or very weak for the proteins with a truncated bHLH domain and for the protein lacking the C-terminal domain which explained the lack of transcriptional activity of these proteins (Neurog31-124, Neurog391-214 and Neurog31-138). The other proteins including the bHLH domain alone were able to bind DNA, and for the bHLH domain it is worth noticing that it gave a stronger band than all other proteins except from Neurog374-214.

The \textit{in vitro} activity of the proteins on the Neurod1 promoter reflected the DNA binding quite well, with the exceptions that in spite of DNA-binding the bHLH domain was not able to induce transcription and the activity of Neurog31-150 was reduced. The protein lacking the whole N-
The terminal domain was able to induce wild type levels of transcription suggesting that the N-terminal domain was not required for transcriptional activity. Truncations from the C-terminal end gradually decreased the transcriptional activity of the proteins, and when the whole C-terminal domain was lost the protein did not have any transcriptional activity \textit{in vitro}. The inability to induce transcription in 3T3 cells may be a reflection of tissue specific differences. The P19 cell line of teratocarcinomal origin whereas the 3T3 cell line is fibroblastic. A difference between the cell lines in e.g. the presence of E-proteins or tissue-specific co-factors may cause such discrepancies, although to my knowledge none have been reported. Immune stainings or pcr of the cell lines would have shown if they differed in the content of E-proteins.

The ability of the truncated proteins to induce differentiation and migration \textit{in vivo} reflected the findings from the luciferase assay for most of the proteins. However, the protein lacking the C-terminal domain was active in two out of six embryos although the luciferase assay suggested that this protein was inactive and the EMSA had suggested very poor DNA binding of this protein. This may reflect that the \textit{in ovo} endoderm electroporation studies are more sensitive than luciferase assays and EMSAs. If the protein has a transcriptional activity level too low to be detected in the luciferase assay it might still be high enough to induce a response \textit{in vivo}. The low activity of Neurog3\textsubscript{1-138} may be close to a threshold level in this assay too, since differentiation and migration is not induced in all embryos. This would make the function of the protein dependent of experimental circumstances such as electroporation efficiency and the number of expression vectors taken up by each cell. However, the \textit{in vitro} and \textit{in vivo} studies are performed with two different cell systems and other differences such as differentiation state, chromatin structure, or cell specific expression of other proteins may influence the response as well. Neurog1\textsubscript{1-150} that showed reduced activity in the luciferase assay was able to induce differentiation and migration of the electroporated cells as well as the wild type protein \textit{in vivo}, based on visual appearance of the embryos. The bHLH domain alone was not sufficient to induce transcription \textit{in vitro} or differentiation and migration \textit{in vivo}.

The luciferase assay suggested that the N-terminal domain of Neurog3 was not required for full transcriptional activity. This was supported by the \textit{in ovo} electroporations, that suggested that most of the C-terminal domain (amino acid 150-214) was likewise not required for full Neurog3 activity. However, the region just C-terminal to the bHLH domain was required for full transcriptional activity (amino acid 138-150). For the Nkx2-2 and Pax4 promoters the bHLH domain was likewise not sufficient to induce transcription (Smith et al., 2003; Watada et al., 2003). In these studies the bHLH domain is defined not to contain the amino acids 132-138, which may also influence the transcriptional activity of the proteins (Smith et al., 2003; Watada et al., 2003). For the ability of Neurog3 to induce transcription from the Pax4
promoter the bHLH domain and the C-terminal domain was required like we report for the Neurod1 promoter, whereas for the Nkx2-2 promoter the bHLH domain and either the C- or the N-terminal domain was required (Smith et al., 2003; Watada et al., 2003). This may suggest that in combination with the bHLH domain, regions in other parts of the protein are required to induce transcription of different promoters.

Although the *in ovo* assay is very sensitive, the variability between embryos electroporated with the same expression vector was quite pronounced with regard to the number of electroporated cells expressing hormones, the number of cells migrating and how far they migrated, and how much they tend to cluster. This great variability between individual embryos may account for the fact that no difference with regard to glucagon-inducing potential was observed between any of the active proteins. The choice to evaluate the potential of the proteins to induce differentiation as the percentage of electroporated cells expressing glucagon may be too narrow. Within 48 hours of transfection, wild type Neurog3 only induces glucagon expression in cells and a very few somatostatin expressing cells (described in the manuscript included as appendix 1). A better measure of differentiation maybe would have been to stain the embryos for all endocrine hormones which would have taken into account the possibility that some of the truncated proteins would maybe induce differentiation of the other endocrine cell types within the 48 hours (β-, δ-, PP, or ε), as it is suggested for Neurog3^{74-214} to induce insulin expression.

It was quite surprising that insulin expression was observed in three out of seven embryos electroporated with the protein lacking the N-terminal domain, especially since it was observed in the duodenum that usually does not express insulin (Neurog3^{74-214}). As mentioned, a study described in the manuscript enclosed as appendix 1 showed that a longer incubation time of the embryos after electroporation would allow for PP and somatostatin to be expressed in Neurog3 electroporated cells, but insulin expression was never induced. How Neurog3^{74-214} induced insulin expression is unknown and it would be highly interesting to investigate this further. Perhaps it is a result of the very good DNA binding abilities of this protein, although this was not reflected by a change in transcriptional activity in the luciferase assay. Alternatively, a change in the co-factors interacting with the protein may have occurred. Another possibility is that the preference of this protein to the different promoters it activates may have changed so that a promoter that was activated to a low degree by the wild type protein may be activated to a higher degree by the truncated protein or vice versa. The luciferase assay however did not suggest a change in transcription activity of the protein from the Neurod1 promoter. To investigate further if Neurog3 would be capable of inducing cell
specification, it would be relevant also to investigate if expression of the other endocrine hormones would be induced by Neurog3\textsuperscript{24-214} or any of the other truncated proteins.

If the finding that the protein lacking the N-terminal domain is able to induce differentiation of insulin expressing cells can be confirmed by repetitions of this experiment, it would suggest that Neurog3 itself would be able to act not only as an inducer of endocrine fate but also to direct the endocrine progenitors to their specific lineage fate. Neurog2 was previously demonstrated to induce different lineages specification of neuronal progenitors into a motor neuron or dendritic fate depending upon the phosphorylation status of the protein (Hand et al., 2005; Ma et al., 2008). Perhaps a similar mechanism could apply to Neurog3 in endocrine differentiation.

Interestingly, the activation domain previously mapped to amino acid 190-214 was not required for full transcriptional activity in the luciferase assay using the Neurod1 promoter or for inducing differentiation in the \textit{in vivo} assay in this report (Neurog3\textsuperscript{1-178}). The previous reported transcriptional domain was mapped using Gal4 fusions, and the discrepancy of these results and the findings of this report stress the importance of confirming data by different approaches and of supplementing \textit{in vitro} results with \textit{in vivo} experiments (Smith et al., 2004).

\subsection*{8.2 The role of E12 in Neurog3 function}

Co-transfecting the truncated proteins with E12 did not affect the cellular localisation except for one of the proteins. The protein that had a C-terminally truncated bHLH domain (Neurog\textsuperscript{1-124}) was present in cytoplasm to a larger degree when transfected alone, but in the presence of exogenous E12, Neurog\textsuperscript{1-124} was primarily located in the nucleus. This effect was different from the observations from the co-transfection studies with Neurog\textsuperscript{1-124} and wild type Neurog3, where both proteins were more equally distributed between the nucleus and the cytoplasm. It is not known if the observation that co-transfected E12 tend to locate the truncated protein to the nucleus is a direct effect or an indirect effect, but as Neurog3 is known to dimerize with E12 it would be obvious to suggest that nuclear localisation domains of E12 would locate the Neurog3/E12 dimer to the nucleus (Lassar et al., 1991; Sun and Baltimore, 1991). This effect of E12 is not as clear for the proteins lacking the N-terminal domain or with a N-terminally truncated bHLH domain, as they are still present in the cytoplasm to large degree, although the nuclei also stain for the truncated proteins. This difference may again be related to the sequence flanking the bHLH on the N-terminal side, and it supports the hypothesis that the
amino acids 74-91 and amino acids within the region of amino acid 56-74 are involved in nuclear localisation of Neurog3.

When the truncated proteins or wild type protein are co-electroporated with E12 into early gut tube chicken endoderm it has a quite remarkable effect. When electroporated with wild type protein or one of the active truncated proteins the electroporated cells migrate but fail to cluster and differentiate into glucagon expressing cells. However, the fact that E12 did not increase the differentiation potential of Neurog3 in vivo in this report was supported by results from the luciferase assay that showed that exogenous E12 did not increase the in vitro activity of Neurog3. When exogenous E12 was added in a 1:3 or 1:1 ratio (E12 to Neurog3) it rather tended to reduce Neurog3 activity although this was not statistically significant. This was surprising taking into account that E12 was required for DNA binding of Neurog3 in EMSA assays and the fact that E47 was previously reported to increase Neurog3 transcriptional activity in luciferase assays on the Neurod1 and Nkx2-2 promoter (Huang et al., 2000; Watada et al., 2003). An E12 mediated increase of Neurog3 half-life would also have been expected to increase Neurog3 activity (Vosper et al., 2007). In ovo electroporations with E12 alone or in combination with any of the in vivo inactive truncated proteins showed that E12 induced clustering of the electroporated cells by itself. Since only a very limited number of electroporated cells is observed outside the endoderm it is speculated that this is an effect of cell proliferation, although it may also be a result of migration within the endoderm. Staining of the cells with BrdU or MPM-2 has previously shown that Neurog3 and E12 co-electroporated cells were proliferative (Yassin, 2008).

The mechanism behind the effect of exogenous E12 on Neurog3 function in vivo is not known. However, since a Neurog3 independent effect of E12 is seen in the electroporated cells it is speculated that the reduced differentiation and clustering may result from these other effects of E12 in the cells and not from a direct E12 mediated inhibitory effect on Neurog3 function. It is remarkable however, that the migratory Neurog3 effect is not reduced by E12 when the differentiation and clustering effects are. However, the differentiation and migratory effects have previously been demonstrated to be uncoupled, as Notch1 mediated inhibition of Neurog3 function only inhibited the differentiation but not migration of the electrporated cells (Ahnfelt-Ronne et al., 2007a). It was previously suggested that the bHLH proteins may induce distinct effects in cells depending e.g. on phosphorylation of the protein. Phosphorylation of Neurog2 specified the migration properties and the dendritic morphology of pyramidal neurons independent of DNA-binding whereas the phosphorylation was not required for transcriptional activity or pro-neural functions (Hand et al., 2005). Also reported for Neurog2, phosphorylation of two other amino acids would increase interaction with a specific
transcription factor and induce motor neuron specification as opposed to neurogenesis, where the phosphorylations was not required (Ma et al., 2008). A similar mechanism may apply to Neurog3 to separate the differentiation and migration functions. Interaction of Neurog3 with exogenous E12 or Neurog3-independent effects of E12 may influence one such modification or protein interaction, although this is pure speculation. Perhaps it might also influence other putative protein-protein interactions, like the ones reported for Neurog3 and Foxa2 or Hnf1a (Smith et al., 2003; Watada et al., 2003). It would have been interesting to investigate the presence of endogenous E-proteins in the chicken gut endoderm by q-pcr, or to knock down the endogenous E-proteins to investigate if this would have an effect on Neurog3 function. A specific knock-down of E12 or E47 would maybe reveal which endogenous E-protein Neurog3 would interact with to mediate its \textit{in vivo} function. A hypothetical phenotype in knock-down experiments may be hard to interpret based on the fact that exogenous E12 electroporated alone had effects in the chicken endoderm.

8.3 Studies of the Neurog3 mutants discovered in humans

The human Neurog3 mutants were shown to have a reduced transcriptional activity on the Neurod1 promoter compared to the wild type protein but retained the ability to induce clustering and glucagon expression \textit{in ovo} electroporation studies (Jensen et al., 2007). The Neurog3 construct containing both mutations had reduced transcriptional activity in a luciferase assay and was inactive \textit{in vivo} (Jensen et al., 2007). In this report, the EMSA showed that the single mutants bound to DNA whereas the double mutant did not or too weakly to be detected. E12 co-electroporated with the double mutant or Neurog3\textsuperscript{R107S} showed the same effect as for the truncated proteins; migration with reduced differentiation and clustering for the single mutant and clustering of GFP expressing cells in the endoderm electroporated with the inactive protein and E12 alone.
The combination of the luciferase assay, \textit{in ovo} electroporation and EMSA provides a good tool for evaluating the impact of mutations on the function of a transcription factor like Neurog3.

8.4 Neurog3 as an activator of Myt1 transcription

The Myt1 6- and 7-zinc finger forms were able to interact but located to different cellular compartments, the 6-zinc finger form primarily to the cytoplasm and the 7-zinc finger form primarily to the nucleus. These findings were based on overexpression, and perhaps an interaction would not occur under normal circumstances. Whether an interaction would have a
functional significance is unknown. It is also not known if the two forms of the protein have the same function. Neurog3 was able to induce transcription of the Myt1 promoter although the induction level was lower for the Myt1 promoter than for the Neurod1 promoter. Again the response induced in 3T3 cells were much lower than the response induced in P19 cells. The reason for this is uncertain but as mentioned earlier, perhaps unidentified cell specific differences may account for it, e.g. differential expression of general transcriptional co-factors or protein interaction partners. The luciferase assay showed a slightly different response of the reporter to the truncated proteins than was observed for the Neurod1 reporter. On the Myt1 promoter the Neurog3$^{589\text{D}}$ negative control induced a slight response. In contrast to the observations on the Neurod1 promoter, the protein that lacked the N-terminal domain (Neurog3$^{74-214}$) and the protein lacking the most C-terminal domain (Neurog3$^{1-178}$) induced a response that was significantly lower than the wild type protein. Hence, the loss of the previously reported activation domain reduced the transcriptional activity on the Myt1 promoter but not on the Neurod1 promoter (Smith et al., 2004). However, the loss of the activation domain did not completely abolish Neurog3 transcriptional activity.

Like on the Neurod1 promoter, the addition of exogenous E12 did not enhance Neurog3 mediated transcriptional activity from the Myt1 promoter, and the same tendency of E12 to reduce the transcriptional activity was observed, although it was not significant.

The differences in Neurog3 activation of the Neurod1 and the Myt1 promoter are small and may have several reasons. First of all, the induction level of the reporters may not reflect the transcription activity from the native gene. Perhaps the differences in transcriptional activity are due to incomplete promoter sequences, maybe some enhancer sequences or other regulatory domains are missing in the reporters. If this is not the case, transcriptional activity on the Myt1 promoter may be lower as a result of lower affinity of Neurog3 to this promoter, or perhaps post-translational modifications or protein interactions would increase Neurog3 activity on this promoter. Truncations from both ends of the protein had a greater effect on the activity from the Myt1 promoter. Perhaps this reflects a lower affinity of Neurog3 to the Myt1 promoter or the requirement of interaction partners to obtain larger responses (Wang et al., 2008). The fact that Neurog3 is not the only inducer of Myt1 expression supports this latter hypothesis, although it has not been investigated (Wang et al., 2008). However, in spite of the minor differences between the response to the truncated Neurog proteins the overall profile is quite similar, suggesting that they are both direct targets of Neurog3.
8.5 Conclusions

It is well known that Neurog3 induces endocrine differentiation and is required for development of the entire endocrine compartment of the pancreas. However, the immediate events following Neurog3 expression is not very well understood. The findings in this thesis provide support to the previously reported observations that Neurog3-induced differentiation and migration are independent processes; that Neurog3 induce transcription from several target genes; and that E12 unexpectedly did not increase Neurog3 *in vitro* activity and did not enhance Neurog3 differentiation potential *in vivo*. We show that the transcription factor Myt1 is a direct target of Neurog3. An intact bHLH domain and amino acids in the region flanking the bHLH domain on the N-terminal side is involved in nuclear localisation. The finding that a N-terminally truncated version of Neurog3 was capable of inducing ectopic insulin expression suggests that Neurog3 may have a function in specification of the endocrine cell types in addition to the function as an inducer of a general endocrine fate, although further investigations are needed to support this hypothesis. A more detailed understanding of Neurog3 regulation and function may provide knowledge that can be applied to the directed differentiation of stem cells into insulin producing cells that may be suited for transplantation.
9. References


Sommer L, Ma Q, Anderson DJ. 1996. Neurogenins, a novel family of atonal-related bHLH transcription factors, are putative mammalian neuronal determination genes that reveal progenitor cell heterogeneity in the developing CNS and PNS. Mol Cell Neurosci 8:221-241.


