Glucocerebrosidase Expression and Analysis

by

Tessa Nicole Campbell
B.Sc., University of Victoria, 1998

A Dissertation Submitted in Partial Fulfillment of the
Requirements for the Degree of

DOCTOR OF PHILOSOPHY

In the Department of Biology

Supervisor: Dr. Francis Y.M. Choy

ABSTRACT

Gaucher disease, an autosomal recessive disorder, is characterized by a heterogeneous set of signs and symptoms caused by a deficiency in the lysosomal enzyme glucocerebrosidase. As a single gene enzyme deficiency, Gaucher disease is a prime candidate for enzyme replacement therapy. Such therapy exists, though the exorbitant cost prevents many from receiving treatment. Thus, a more cost-effective method of producing glucocerebrosidase was examined. The *Pichia pastoris* yeast system was chosen, but resultant production levels were low. Two variants of green fluorescent protein (GFP), red-shifted GFP (RSGFP) and enhanced GFP (EGFP), were employed as molecular reporters to track enzyme production and isolation. No expression of glucocerebrosidase was evident, indicating that the *P. pastoris* system was not an appropriate choice for glucocerebrosidase production. Both GFP variants were successfully expressed, with EGFP levels apparently greater than RSGFP levels. To study glucocerebrosidase production and trafficking in a higher eukaryotic system, EGFP-tagged glucocerebrosidase constructs were expressed in HeLa cells. Though EGFP was readily visualized, few cells expressing glucocerebrosidase constructs were present. No co-localization with organelle markers was evident. Examination at the RNA level indicated successful transcription, however, an apparent translational inefficiency was encountered. To shed light on the possible cause of this inefficiency, two approaches were taken: one examined expression of truncated glucocerebrosidase constructs in HeLa cells, the other included co-transfection with small interfering RNAs (siRNAs) in both HeLa and COS-1 cells. In the first approach, greater expression was seen from the EGFP-tagged construct devoid of the proposed inhibitory binding site than from the EGFP-tagged construct containing the binding site. Expression of both truncated constructs was greater than that of EGFP-tagged glucocerebrosidase starting at either initiation codon, indicating a more complex mechanism of translational control than strictly inhibition from the proposed site. In the second approach, a siRNA was designed to block TCP80, which has been suggested to inhibit glucocerebrosidase translation. Co-transfection

studies of siRNAs (control, EGFP and TCP80) and glucocerebrosidase/ EGFP plasmids were performed in HeLa and COS-1 cells. In both cell types, all constructs were successfully expressed when co-transfected with control siRNA, as indicated by RNA and protein examination. Introduction of TCP80 siRNA in both cell types did not serve to increase glucocerebrosidase expression as expected, but instead decreased such expression. EGFP expression was readily knocked down in HeLa and COS-1 cells by GFP-targeted siRNA. Knockdown was evident in the expression of glucocerebrosidase/EGFP constructs, indicating that fusion with EGFP may serve as a means to introduce a foreign gene, then knock its expression down at a desired time by introduction of a GFP-targeted siRNA.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGMENTS

I would like to thank my Supervisor, Dr. Francis Choy, for his constant encouragement and tissue culture assistance over the years. I would also like to extend my gratitude to all my labmates, Chris Lamb, the Burke lab, the Hintz lab, the Koop lab, the Advanced Imaging lab, the Biology Office staff, the Graduate Studies Office staff, the Natural Sciences and Engineering Research Council, the Michael Smith Foundation for Health Research, and the Scottish Rite Charitable Foundation of Canada. To my family and friends, thank you for your support. To Eileen Campbell, Wayne Campbell, Sean Campbell, and Stacey Olesky, thanks for always being there. Finally, to my future husband, Chris Davidson, thank you for your unconditional love and for understanding that research never ends.

Chapter 1 - Gaucher Disease

1.1 History and Clinical Manifestations

In 1882, Dr. Philippe Charles Ernest Gaucher first described a neurologically normal female with massive hepatosplenomegaly. He believed her disease to be a primary splenic neoplasm (Beutler and Grabowski 2001). By 1905, the systematic and familial nature of the disease was recognized, leading to the coining of the phrase "Gaucher disease" by Brill. Though the metabolic nature of the disease was determined by Marchand in 1907, it was not until 1934 that the major accumulated lipid was identified as glucocerebroside (glucosylceramide) by Aghion (Zhao and Grabowski 2002, Beutler and Grabowski 2001). In 1965, both Patrick and Brady et. al. identified impaired glucocerebroside hydrolysis to be the enzymatic defect in Gaucher disease (Brady et. al. 1965, Patrick 1965).

Gaucher disease is an autosomal recessive disorder characterized by a heterogeneous set of signs and symptoms caused by the defective lysosomal hydrolysis of glucocerebroside and related glucosphingolipids. This defective hydrolysis results from dysfunction of the enzyme glucocerebrosidase (glucosylceramidase, acid ß-glucosidase, EC 3.2.1.45) which is involved in the conversion of glucocerebroside to glucose and ceramide (Beutler and Grabowski 2001). Accumulated glucocerebroside is taken-up by macrophages and subsequently deposited in the spleen, liver, and bone marrow. Histological examination of affected tissues reveals lipid engorged macrophages (Gaucher Cells) which lead to the disruption of normal organ function. The resultant hepatosplenomegaly, bone crises, and pancytopenia are the characteristic symptoms of

the disease. Gaucher disease is currently the most common lysosomal storage disorder and the first to be successfully treated by enzyme replacement therapy (de Fost et. al. 2003). Three major types of the disease have been classified based on the absence or presence and severity of primary central nervous system involvement: Type 1 (nonneuronopathic), Type 2 (acute neuronopathic), and Type 3 (subacute neuronopathic) (de Fost et. al. 2003).

Type 1 Gaucher disease, the mildest and most common form, is characterized by the lack of neuronopathic involvement. The mean age at diagnosis is 21 years, but age of onset can range from early childhood to the eighth decade (de Fost et. al. 2003). The variability of visceral manifestation ranges from fatal disease in the first two decades of life to essentially asymptomatic octogenarians (Zhao and Grabowski 2002). Although patients vary widely in levels of visceral and skeletal involvement, splenic enlargement is present in all symptomatic patients (Beutler and Grabowski 2001).

Type 2 Gaucher disease, the most severe form, is characterized by severe early-onset neuronopathy and early mortality. Gaucher cells have been shown to accumulate in the sub-cortical white matter, but the cause of neuronal loss remains to be elucidated (Balicki and Beutler 1995). Time of onset varies from birth to 6 months, with death of most infants occurring by two years of age. Extensive visceral involvement with hepatosplenomegaly is the hallmark phenotype. Oculomotor abnormalities, limb rigidity, and seizures are also common (Grabowski 1993, Beutler 1995).

The third form of Gaucher disease, Type 3, is characterized by sub-acute neuronopathic involvement: patients show neurodegenerative symptoms but are able to survive through childhood into adulthood. Massive visceral involvement is usually

present. The prototype of Type 3 is the Norrbottnian form, with a median age of onset of visceral symptoms at one year. Similar to Type 2, the first neurologic symptoms are generally disorders of eye movement within the first decade of life (Beutler and Grabowski 2001).

1.2 Molecular Genetics

Both the full length cDNA (Sorge et. al. 1985) and genomic DNA (Horowitz et. al. 1989) sequences of glucocerebrosidase have been elucidated. The 7.6 kilobase (kb) gene, which resides at 1q21 (Shafit-Zagardo et. al. 1981, Ginns et. al. 1985), is composed of 11 exons and possesses TATA and CAAT boxes in the promoter region (Reiner et. al. 1988, Horowitz et. al. 1989). Two in-frame ATG sites are present, both of which appear to be functional *in vitro* and *in vivo* (Sorge et. al. 1987, Pasmanik-Chor et. al. 1996). Translation from the first ATG results in a glucocerebrosidase peptide harbouring a leader of 39 amino acids: the first 20 comprise a highly hydrophilic sequence, while the latter 19 represent a hydrophobic sequence. Translation from the second ATG results in a peptide with a leader of 19 hydrophobic amino acids (Sorge et. al. 1987).

Approximately 16 kb downstream of the glucocerebrosidase gene is a 5.8 kb pseudogene with 96% sequence similarity. The pseudogene contains deletions in exon 9, as well as in introns 2, 4, 6 and 7. Additionally, numerous base pair changes are present throughout the gene (Horowitz et. al 1989, Zimran et. al. 1990). Although it is transcribed, the pseudogene is not successfully translated due to incorrect splicing and absence of a long open-reading frame (Sorge et. al. 1990).

Immediately downstream from the functional glucocerebrosidase gene is the pseudometaxin gene, while immediately downstream from the glucocerebrosidase pseudogene is the metaxin gene (Bornstein et. al. 1995, Long et. al. 1996). Though contiguous with the glucocerebrosidase gene and pseudogene, the metaxin gene and pseudogene are in a reverse orientation. Apparently, a large block of DNA containing the ancestral glucocerebrosidase and metaxin genes underwent duplication, giving rise to the two corresponding pseudogenes in the lineage leading up to modern *Homo sapiens*. This duplication event is present in the Great Apes and Old World Monkeys and has been estimated to have occurred 36-40 million years ago (Winfield et. al. 1997).

1.3 Mutations

Mutations in the glucocerebrosidase gene will be represented in this thesis by the resultant amino acid substitution (i.e. N370S as a substitution of serine for asparagine at amino acid 370 of the mature polypeptide). More than 200 missense, termination, rearrangement, deletion, and insertion mutations at the glucocerebrosidase locus have been identified in Gaucher patients (Grabowski 2003). Most of these are rare and/or private mutations, however, several have significant frequencies. Two such mutations, L444P in exon 10 and N370S in exon 9, account for approximately 60% of all alleles present in Gaucher cases (Grabowski 1993). Homozygosity for the L444P mutation is usually associated with neuronopathic disease. In contrast, both homozygosity and heterozygosity for the N370S mutation are associated with the Type 1 (noneuronopathic) form (Beutler and Grabowski 2001).

Although Gaucher disease is a rare panethnic disorder, disease frequency is elevated in the Ashkenazi Jewish population: 1/500 live births versus 1/60,000-1/360,000 in the white non-Jewish population (Colombo 2000). Mutations N370S, L444P, ins84GG (insertion mutation causing a frameshift), and IVS2 (+1) (splice donor site variant in intron 2) account for 93% of the mutations identified in Ashkenazi Jewish Type 1 patients, while these same four mutations only account for 49% of mutant alleles in the non-Jewish population (Koprivica et. al. 2000). An explanation for such increased disease allele frequencies within this population has not been reached, but theories of founder effects and possible heterozygous selective advantage have been proposed (Diamond 1994, Colombo 2000, Diaz et. al. 2000).

1.4 Biochemistry and Cell Biology

Glucocerebrosidase, a lysosomal membrane-associated hydrolase, has been purified from a number of species, but the human form has been the focus of the majority of investigations (Beutler and Grabowski 2001). The mature protein is 497 amino acids, excluding the 19 or 39 amino acid leader sequence that is cleaved during transit through the endoplasmic reticulum (Sorge et. al. 1985). Seven cysteines are present at residues C4, C16, C23, C128, C248, and C342, with the first four participating in disulfide formation (Zhao and Grabowski 2002). Though no transmembrane domains or large tracts of hydrophobicity have been predicted, detergents are required to solubilize the protein from the lysosomal membrane prior to purification (Erickson et. al. 1985, Choy and Woo 1991, Qi and Grabowski 2001).

The natural substrates for glucocerebrosidase are glucosphingolipids with medium to long fatty-acyl chains. The identity of the substrate sugar moiety appears to be of greater importance than the aglycan portion of the molecule (Glew et. al. 1988). Consequently, an *in vitro* fluorogenic assay using 4-methyl-umbelliferyl-ß-D-glucopyranoside (4MUGP) as an alternate substrate has been developed to measure glucocerebrosidase activity (Daniels et. al. 1980). The protein, with the active site located near the carboxy terminus, is optimally functional at pH 5.5 (Glew et. al. 1988, Beutler and Grabowski 2001) and requires glycosylation at the first sequon for activity (Berg-Fussman et. al. 1993). Glucocerebrosidase also apparently requires association with negatively charged phospholipids and Saposin C for activity. Though the mechanism of activation is still under debate, the following theory has been proposed: upon association with the lysosomal membrane, glucocerebrosidase attaches to negatively charged phospholipids and undergoes a conformational change which realigns active site residues. This step, though slowly reversible, now permits catalytic activity. Saposin C attaches to the glucocerebrosidase/phospholipid complex and induces a further conformational change, leading to a fully active enzyme (Zhao and Grabowski 2002).

Glucocerebrosidase biosynthesis has been examined in cultured porcine kidney cells and human skin fibroblasts (Erickson et. al. 1985, Beutler and Kuhl 1986, Bergman and Grabowski 1989). Molecular weight ranges from 58,000 to 66,000 depending upon the amount of glycosylation (Balicki and Beutler 1995). Cotranslational glycosylation at four of the five sequons and signal peptide cleavage occur as glucocerebrosidase moves through the endoplasmic reticulum membrane (Erickson et. al. 1985, Berg-Fussman et. al. 1993). Within 72 hours, the initial high mannosyl chains are transformed into complex

types and remodelled within the Golgi apparatus to yield the final mature protein (Takasaki et. al. 1984, Erickson et. al. 1985). Total deglycosylation yields a protein with a molecular weight of 55,000 to 56,000 (Sorge et. al. 1985).

Transport from the Golgi to the lysosomes does not follow the typical mannose-6-phosphate receptor pathway of most lysosomal hydrolases. Evidence for this exception comes from examination of the lysosomes of patients with I-Cell disease, an inherited deficiency in phosphotransferase which results in the absence of mannose phosphorylation (Reitman et. al. 1981). Glucocerebrosidase is found at normal levels in I-Cell patient lysosomes, whereas most other lysosomal hydrolases are mistargeted and secreted (Leroy et. al. 1972, Wenger et. al. 1976, Lemansky et. al. 1985).

Glucocerebrosidase transport to the lysosome also does not appear to utilize a cytoplasmic tail signal. Acid phosphatase, for example, is rapidly endocytosed from the cell surface due to a tyrosine-containing internalization signal in its 19 amino acid cytoplasmic tail (Lehmann et. al. 1992). Computer analysis of the glucocerebrosidase core, however, fails to identify any transmembrane domain aside from the signal peptide that is cleaved following transport into the endoplasmic reticulum (Beutler and Grabowski 2001).

The role of glycosylation has been examined with respect to glucocerebrosidase trafficking. Some evidence has suggested a role for glycosylation in lysosomal targeting (Aerts et. al. 1986), while other evidence suggests no role at all (Leonova and Grabowski 2000). In the presence of glycan formation inhibitors, glucocerebrosidase has been demonstrated to become membrane-associated (Rijnboutt et. al. 1991).

Since transport to the lysosome does not appear to be directed by mannose phosphorylation, a cytoplasmic tail, or glycosylation, an alternative mechanism must be at work. One possibility is that association with other proteins may play a role. Elevated levels of lysosomal-associated membrane proteins (LAMPs), Saposin C, and cathepsin D have been observed in Gaucher patients' lysosomes (Zimmer et. al. 1999). Thus, a protein-protein or glycan-protein association with these other species may be involved in glucocerebrosidase transport to the lysosome.

1.5 Therapeutic Strategies

Three main therapeutic approaches have been presented for Gaucher disease: bone marrow transplantation, enzyme replacement therapy, and gene therapy. Most of the Gaucher symptoms result from lipid accumulation in macrophages, which are progeny of hematopoietic stem cells. Thus, it follows that allogeneic bone marrow transplantation has been one of the therapeutic strategies developed for treatment of Gaucher patients. This approach resulted in the first effective treatment for Gaucher disease (Rappeport and Ginns 1984) and has been employed in a number of other cases (Beutler and Grabowski 2001). Such treatment has resulted in the disappearance of hepatomegaly and a decrease in growth delay and bone pain (Hoogerbrugge et. al. 1995). Although the response to transplantation has been positive in surviving patients, there have been other deaths secondary to the transplantation procedure (Rappeport et. al. 1984, Hoogerbrugge et. al. 1995, Beutler and Grabowski 2001). In one study of 63 transplant recipients spanning 14 lysosomal storage disorders, there was a 10% mortality rate if an HLA identical sibling marrow donor was present and a 20-25% mortality rate if mismatched tissue was

available (Hoogerbrugge et. al. 1995). Due to the high risk associated with transplantation, it is not surprising that enzyme replacement therapy remains a more popular choice for Gaucher patients.

Enzyme replacement therapy for Gaucher disease was first proposed in 1966 (Brady 1966). In 1974, it was demonstrated that single intravenous infusions of purified placental glucocerebrosidase reduced hepatic and blood glucocerebroside levels (Brady et. al. 1974). Based on these initial results, a large-scale purification method was developed for glucocerebrosidase (Furbish et. al. 1977). Early trials proved disappointing, however, because most of the enzyme was mistargeted (Barton et. al. 1991). Several strategies that took advantage of the mannose lectin on the macrophage plasma membrane were examined, with sequential deglycosylation to expose inner mannose residues on the oligosaccharide chains of glucocerebrosidase being the most effective (Barton et. al. 1991). Based on this improvement in targeting, two commercial enzymes have been developed to become the standard in Gaucher disease treatment: aglucerase (*Ceredase*™) and imiglucerase (*Cerezyme*™). Aglucerase is modified human glucocerebrosidase purified from pooled human placentae, while imiglucerase is produced from human glucocerebrosidase cDNA in Chinese Hamster Ovary (CHO) cells (Grabowski et. al. 1998). Both commercial enzymes have been effective in reducing organomegaly, bone pain, and haematological complications (Beutler and Grabowski 2001). The enzymes have not been effective, however, in altering the progression of neurodegeneration in Type 2 and Type 3 patients. Moreover, the enormous cost of treatment, $60-$240 USD/kg body weight every 2 weeks (translating to >$100,000-$400,000 per 70 kg patient per year), limits its availability to many Gaucher patients

(Grabowski et. al. 1998). Thus, other treatment avenues, such as gene therapy, are currently being pursued.

Since gene therapy is a potentially curable treatment, considerable effort has been expended to develop efficient gene transfer. Based on the successes of bone marrow transplantation whereby a portion of a patient's defective hematopoietic stem cells are replaced with a donor's normal stem cells, it follows that correction of the defect in the patient's own cells should be an effective method of treatment. However, because the "corrected" stem cells would not have a proliferative advantage over the defective cells, cure would be expected only if the patient's defective cells were at least partially ablated. This risk is compounded by the difficulty of the transfer of functional glucocerebrosidase genes into hematopoietic stem cells (Grabowski 1993, Beutler and Grabowski 2001, Grabowski 2003). Another current promising approach, however, involves the transduction of isolated myoblasts (muscle cells) with a retroviral vector containing human glucocerebrosidase cDNA (Liu et. al. 1998). Transduced murine and human myoblasts had intracellular glucocerebrosidase activities 5-10 times those of non-transduced controls and continued to secrete the enzyme for up to 35 weeks *in vitro* (Liu et. al. 1998). Nevertheless, no confirmed successful clinical trials for gene therapy have been demonstrated to date.

1.6 Dissertation Outline

Chapter 1 of this dissertation addresses the backgrounds of Gaucher disease and glucocerebrosidase, thus providing a framework for the following chapters. In Chapter 2, the production and isolation of two green fluorescent protein (GFP) variants and GFP-

tagged glucocerebrosidase in *Pichia pastoris* are examined to shed light upon enzyme production for therapeutic purposes. In Chapter 3, glucocerebrosidase is expressed in a mammalian system (HeLa cells) to provide insight into *in vivo* enzyme biosynthesis and trafficking in a higher eukaryotic system. Chapter 4 utilizes two approaches to build upon the information regarding translational inefficiency gleaned from the previous chapter. The first approach employs truncated GFP-tagged glucocerebrosidase constructs, while the second takes advantage of a newly reported phenomenon termed RNA interference. Finally, Chapter 5 provides a summary of the results of each previous chapter and offers some future perspectives based upon the discoveries that have arisen from the research reported in this dissertation.

1.7 References

Aerts, J.M.F.G., Brul, S., Donker-Koopman, W.E., van Weely, S., Murray, G.J., Barranger, J.A., Tager, J.M., and Schram, A.W. (1986). Efficient routing of glucocerebrosidase to lysosomes requires complex oligosaccharide chain formation. *Biochem. Biophys. Res. Comm.* 141: 452-458.

Balicki, D. and Beutler, E. (1995). Gaucher disease. *Rev. Mol. Med.* 74: 305- 323.

Barranger, J.A., Tomich, J., Weiler, S., Sakallah, S., Sansieri,C., Mifflin, T., Bahnson, A., Wei, F-S., Wei, J-F., Vallor, M., Nimgaonkar, M., Ball, E., Mohney, T., Dunigan, J., Ohashi, T., Bansal, V., Mannion-Henderson, J., Liu, C., and Rice, E. (1995). Molecular biology of glucocerebrosidase and the treatment of Gaucher disease. *Cytokines and Molecular Therapy* 1:149-163.

Barton, N.W., Brady, R.O., Dambrosia, J.M., Di Bisceglie, A.M., Dopplet, S.H., Hill, S.C., Mankin, H.J., Murray, G.J., Parker, R.I., Argoff, C.E., Grewal, R.P., and Yu, K.-T. (1991). Replacement therapy for inherited enzyme deficiency: macrophage-targeted glucocerebrosidase for Gaucher's disease. *N. Engl. J. Med.* 324: 1464-1470.

Berg-Fussman, A., Grace, M.E., Ioannou, Y., and Grabowski, G.A. (1993). Human acid β-glucosidase: *N*-glycosylation site occupancy and the effect of glycosylation on enzyme activity. *J. Biol. Chem.* 268: 14861-14866.

Bergman, J.E. and Grabowski, G.A. (1989). Posttranslational processing of human lysosomal acid ß-glucosidase: a continuum of defects in Gaucher disease Type 1 and Type 2 fibroblasts. *Am. J. Hum. Genet.* 44: 741-750.

Beutler, E. (1995). Gaucher disease. *Advances in Genetics* 32: 17-49.

Beutler, E. and Grabowski, GA. (2001). Gaucher disease. In Scriver, C.R., Beaudet, A.L, Sly, W.S., and Valle, D (eds.), *The Metabolic and Molecular Bases of Inherited Disease* (vol. 3, 8th ed., pp. 3635-3668). New York: McGraw-Hill.

Beulter, E. and Kuhl, W. (1986). Glucocerebrosidase processing in normal fibroblasts and in fibroblasts from patients with Type I, Type II, and Type III Gaucher disease. *Proc. Natl. Acad. Sci. USA* 83: 7472-7474.

Bornstein, P., McKinney, C.E., LaMarca, M.E., Winfield, S., Shingu, T., Devarayalu, S., Vos, H.L., and Ginns, E.I. (1995). Metaxin, a gene contiguous to both thrombospondin 3 and glucocerebrosidase, is required for embryonic development in the mouse: implications for Gaucher disease. *Proc. Natl. Acad. Sci. USA* 92: 4547-4551.

Brady, R.O. (1966). The sphingolipidoses. *N. Engl. J. Med.* 275: 312-318.

Brady, R.O., Kanfer, J.N., and Shapiro, D. (1965). Metabolism of glucocerebrosides: evidence of an enzymatic deficiency in Gaucher's disease. *Biochem. Biophys. Res. Comm.* 18: 221-225.

Brady, R.O., Pentchev, P.G., Gal, A.E., Hibbert, S.R., and Dekaban, A.S. (1974). Replacement therapy for inherited enzyme deficiency: use of purified glucocerebrosidase in Gaucher's disease. *N. Engl. J. Med.* 291: 989-993.

Choy, F.Y.M. and Woo, M. (1991). Purification and the effect of peptide *N*-glycosidase F on lysosomal membrane-bound glucocerebrosidase from human cultured fibroblasts. *Biochem. Cell Biol.* 69: 551-556.

Colombo, R. (2000). Age estimate of the N370S mutation causing Gaucher disease in Ashkenazi Jews and European populations: a reappraisal of haplotype data. *Am. J. Hum. Genet.* 66: 692-697.

Daniels, L.B., Glew, R.H., Radin, N.S., and Vunnam, R.R. (1980). A revised fluorometric assay for Gaucher's disease using conduritol-ß-epoxide as the source of ß-glucosidase. *Clin. Chim. Acta* 106: 155-163.

de Fost, M., Aerts, J.M.F.G., and Hollak, C.E.M. (2003). Gaucher disease: from fundamental research to effective therapeutic interventions. *Neth. J. Med.* 61: 3-8.

Diamond, J.M. (1994). Jewish lysosomes. *Nature* 368: 291-292.

Diaz, G.A., Gelb, B.D., Risch, N., Nygaard, T.G., Frisch, A., Cohen, I.J., Sa Miranda, C., Amaral, O., Maire, I., Poenaru, L., Caillaud, C., Weizberg, M., Mistry, P., and Desnick, R.J. (2000). Gaucher disease: the origins of the Ashkenazi Jewish N370S and 84GG acid β-glucosidase mutations. *A. J. Hum. Genet.* 66: 1821-1832.

Erickson, A.H., Ginns, E.I., and Barranger, J.A. (1985). Biosynthesis of the lysosomal enzyme glucocerebrosidase. *J. Biol. Chem.* 260: 14319-14324.

Furbish, F.S., Blair, H.E., Shiloach, J., Pentchev, P.G., and Brady, R.O. (1977). Enzyme replacement therapy in Gaucher's disease: large-scale purification of glucocerebrosidase suitable for human administration. *Proc. Natl. Acad. Sci. USA* 74: 3560-3563.

Ginns, E.I., Choudary, P.V., Tsuji, S., Martin, B., Stubblefield, B., Sawyer, J., Hozier, J., and Barranger, J.A. (1985). Gene mapping and leader polypeptide sequence of human glucocerebrosidase: implications for Gaucher disease. *Proc. Natl. Acad. Sci. USA* 82: 7101-7105.

Glew, R.H., Basu, A., La Marco, K.L., and Prence, E.M. (1988). Mammalian glucocerebrosidase: implications for Gaucher's disease. *Lab. Invest.* 58: 5-25.

Grabowski, G.A. (1993). Gaucher disease: enzymology, genetics, and treatment. *Adv. Hum. Genet.* 21: 377-441.

Grabowski, G.A. (2003). Perspectives on gene therapy for lysosomal storage diseases that affect hematopoiesis. *Curr. Hematol. Rep.* 2: 356-362.

Grabowski, G.A, Gatt, S., and Horowitz, M. (1990). Acid β-glucosidase: enzymology and molecular biology of Gaucher disease. *Biochemistry and Molecular Biology* 25: 385-414.

Grabowski, G.A., Leslie, N., and Wenstrup, R. (1998). Enzyme therapy for Gaucher disease: the first 5 years. *Blood Rev.* 12: 115-133.

Hoogerbrugge, P.M., Brouwer, O.F., Bordigoni, P., Ringden, O., Kapaun, P., Ortega, J.J., O'Meara, A., Cornu, G., Souillet, G., Frappaz, D., Blanche, S., and Fischer, A. (1995). Allogeneic bone marrow transplantation for lysosomal storage diseases. *Lancet* 345: 1398-1402.

Horowitz, M., Wilder, S., Horowitz, Z., Reiner, O., Gelbart, T., and Beutler, E. (1989). The human glucocerebrosidase gene and pseudogene: structure and evolution. *Genomics* 4: 87-96.

Koprivica, V., Stone, D.L., Park, J. K., Callahan, M., Frisch, A., Cohen, I.J., Tayebi, N., and Sidransky, E. (2000). Analysis and classification of 304 mutant alleles in patients with Type 1 and Type 3 Gaucher disease. *Am. J. Hum. Genet.* 66: 1777-1786.

Lehmann, L.E., Eberle, W., Krull, S., Prill, V., Schmidt, B., Sander, C., von Figura, K., and Peters, C. (1992). The internalization signal in the cytoplasmic tail of lysosomal acid phosphatase consists of the hexapeptide PGYRHV. *EMBO J.* 11: 4391-4399.

Lemansky, P., Gieselmann, V., Hasilik, A., and von Figura, K. (1985). Synthesis and transport of lysosomal acid phosphatase in normal and I-cell fibroblasts. *J. Biol. Chem.* 260: 9023-9030.

Leonova, T., and Grabowski, G.A. (2000). Fate and sorting of acid ß-glucosidase in transgenic mammalian cells. *Mol. Genet. Metabol.* 70: 281-294.

Leroy, J.G., Ho, M.W., MacBrinn, M.C., Zielke, K., Jacob, J., and O'Brien, J.S. (1972). I-Cell disease: biochemical studies. *Pediatr. Res.* 6: 752-757.

Liu, C., Bahnson, A.B., Dunigan, J.T., Watkins, S.C., and Barranger, J.A. (1998). Long-term expression and secretion of human glucocerebrosidase by primary murine and human myoblasts and differentiated myotubes. *J. Mol. Med.* 76: 773-781.

Long, G.L., Winfield, S., Adolph, K.W., Ginns, E.I., and Bornstein, P. (1996). Structure and organization of the human metaxin gene (*MTX*) and pseudogene. *Genomics* 33: 177-184.

Pasmanik-Chor, M., Elroy-Stein, O., Aerts, H., Agmon, V., Gatt, S., and Horowitz, M. (1996). Overexpression of human glucocerebrosidase containing different-sized leaders. *Biochem. J.* 317: 81-88.

Patrick, A.D. (1965). A deficiency of glucocerebrosidase in Gaucher's disease. *Biochem. J.* 97: 17C-18C.

Qi, X. and Grabowski, G.A. (2001). Molecular and cell biology of acid ß-glucosidase and prosaposin. *Progress Nucleic Acid Res.* 66: 203-239.

Rappeport, J.M., and Ginns, E.I. (1984). Bone-marrow transplantation in severe Gaucher's disease. *N. Engl. J. Med.* 311: 84-88.

Reiner, O., Wigderson, M., and Horowitz, M. (1988). Structural analysis of the human glucocerebrosidase genes. *DNA* 7: 107-116.

Reitman, M.L., Varki, A., and Kornfield, S. (1981). Fibroblasts from patients with I-Cell disease and Pseudo-Hurler Polydystrophy are deficient in uridine 5'-diphosphate-*N*-acetylglucosamine: glycoprotein *N*-acetylglucosaminylphosphotransferase activity. *J. Clin. Invest.* 67:1574-1579.

Rijnboutt, S., Aerts, H.M.F.G., Geuze, H., Tager, J.M., and Strous, G.J. (1991). Mannose 6-phosphate-independent membrane association of cathepsin D, glucocerebrosidase, and sphingolipid-activating protein in HepG2 cells. *J. Biol. Chem.* 266: 4862-4868.

Shafit-Zagardo, B., Devine, E.A., Smith, M., Arredondo-Vega, F., and Desnick, R. (1981). Assignment of the gene for acid β-glucosidase to human chromosome 1. *Am. J. Hum. Genet.* 33 : 564-575.

Sorge, J., Gross, E., West, C., and Beutler, E. (1990). High level transcription of the glucocerebrosidase pseudogene in normal subjects and patients with Gaucher disease. *J. Clin. Invest.* 86: 1137-1141.

Sorge, J.A., West, C., Kuhl, W., Treger, L., and Beutler, E. (1987). The human glucocerebrosidase gene has two functional ATG initiator codons. *Am. J. Hum. Genet.* 41: 1016-1024.

Sorge, J., West, C., Westwood, B., and Beutler, E. (1985). Molecular cloning and nucleotide sequence of human glucocerebrosidase cDNA. *Biochemistry* 82: 7289-7293.

Takasaki, S., Murray, G.J., Furbish, S., Brady, R.O., Barranger, J.A., and Kobata, A. (1984). Structure of the *N*-asparagine-linked oligosaccharide units of human placental ß-glucocerebrosidase. *J. Biol. Chem.* 259: 10112-10117.

Wenger, D.A., Sattler, M., Clark, C., and Wharton, C. (1976). I-Cell disease: activities of lysosomal enzymes toward natural and synthetic substrate. *Life Sci.* 19: 413-420.

Winfield, S.L., Tayebi, N., Martin, B.M., Ginns, E.I., and Sidransky, E. (1997). Identification of three additional genes contiguous to the glucocerebrosidase locus on chromosome 1q21: implications for Gaucher disease. *Genome Research* 7: 1020-1026.

Zhao, H. and Grabowski, G.A. (2002). Human genome and diseases: review . Gaucher disease: perspectives on a prototype lysosomal storage disorder. *Cell. Mol. Life Sci.* 59: 694-707.

Zimmer, K.-P., Le Coutre, P., Aerts, H.M.F.G., Harzer, K., Fukuda, M., O'Brien, J.S., and Naim, H.Y. (1999). Intracellular transport of acid ß-glucosidase and lysosome-associated membrane proteins is affected in Gaucher's disease (G202R mutation). *J. Pathol.* 188: 404-414.

Zimran, A., Sorge, J., Gross, E., Kubitz, M., West, C., and Beutler, E. (1990). A glucocerebrosidase fusion gene in Gaucher disease: implications for the molecular anatomy, pathogenesis, and diagnosis of this disorder. *J. Clin. Invest.* 85: 219-222.

Chapter 2 - Heterologous Expression of Glucocerebrosidase-Green Fluorescent Protein Chimerae in *Pichia pastoris*

2.1 Introduction

2.1.1 Glucocerebrosidase and the Pichia pastoris Heterologous Expression System

In 1991, Genzyme Corporation in cooperation with the National Institute of Health developed aglucerase (*Ceredase*™), a macrophage-targeted glucocerebrosidase isolated from human placenta, for enzyme replacement therapy for Gaucher patients (Genzyme Corporation, 2003). Though effective in reducing organomegaly, bone pain, and haematological complications of Type 1 cases (Beutler and Grabowski 2001), the enzyme required costly labour for its production, resulting in prohibitively expensive treatment. Depending upon the weight of the patient, annual treatment costs could reach upwards of $400,000 USD (Grabowski 1998). As a result, many Gaucher patients could not afford the therapy. Soon after, an effective recombinant enzyme form (imiglucerase, *Cerezyme*™) was produced in a modified CHO cell expression system (Grabowski 1998). Although imiglucerase could be produced at significantly higher levels than aglucerase, the overall cost of the treatment did not correspondingly decrease. In lieu of this discrepancy, the Choy lab began testing another heterologous expression system for production of active glucocerebrosidase. The yeast system *Pichia pastoris* was selected.

The *Pichia pastoris* system was originally developed by Phillips Petroleum Company in the 1970s to generate yeast biomass or single-cell protein to be marketed as high protein animal feed. When the cost of methane increased (the inexpensive source of methanol for yeast induction), the system was not economically competitive and was thus

altered in the early 1980s to become a heterologous protein expression system (Higgins and Cregg 1998). As a eukaryote, *P. pastoris* provides many of the advantages of higher eukaryotic systems such as protein processing, folding, and posttranslational modification, while being faster, easier, and less expensive to use (Higgins and Cregg 1998). Additionally, *P. pastoris* contains a tightly-controlled alcohol oxidase 1 gene (*AOX1*). Expression of *AOX1* is induced by methanol up to 1000-fold, representing 5% of total soluble protein produced in shake-flask cultures, and up to 30% in fermenter cultures. A plasmid-borne version of the *AOX1* promoter can thus be used to drive the expression of the gene of interest encoding the desired protein. To date, over 400 proteins, from human endostatin to spider dragline silk protein, have been produced in *Pichia* (Lin Cereghino and Cregg 2000, Lin Cereghino et. al. 2002). However, despite the successes of other laboratories in the production of numerous proteins, the Choy lab was unable to express significant amounts of glucocerebrosidase. Alterations in growth conditions, plasmids, and cassette copy numbers did not result in desired levels of expression. To further provide insight into optimizing glucocerebrosidase production, another approach was required. The development of green fluorescent protein as a molecular reporter provided such an opportunity to shed light on enzyme production.

2.1.2 Green Fluorescent Protein

It is difficult to peruse a biological, microbiological, or biochemical journal and avoid encountering the phrase "green fluorescent protein". The existence of the green fluorescent protein (GFP) of the jellyfish *Aequorea victoria* was reported decades ago (Shimomura et. al. 1962, Morise et. al. 1974); however, the cloning (Prasher et. al. 1992) and heterologous

expression of its cDNA (Chalfie et. al. 1994) soon ignited an explosion of applications for GFP. Such applications include monitoring of gene expression (Li et. al. 1999, Wheeler et. al. 2000), protein localization (Wang and Hazelrigg 1994, Harada et. al. 2000), host-pathogen interactions (Dhandayuthapani et. al. 1995, Valdivia and Falkow1996), cellular dynamics (Rizzuto et. al. 1995, Gerdes and Kaether 1996), protein purification (Cha et. al. 1999, Dąbrowski et. al. 1999), Ca^{2+} concentration (Miyawaki et. al. 1997, Allen et. al. 1999), and pH levels (Kneen et. al. 1998, Miesenböck et. al. 1998, Campbell and Choy 2001).

Wild-type GFP is 238 amino acids and approximately 27,000 MW (Prasher et. al. 1992). It absorbs maximally at ~393 nm with a minor peak at 473 nm, and emits green light at 509 nm (Ward et. al. 1980). Different mutational approaches have optimized expression by altering the promoter, codon usage, folding, splicing, or ribosome binding. GFP variants with differing spectra have also been created, permitting multicolor microscopic visualization (Rizzuto et. al. 1996, Palm and Wlodawer 1999). GFP fluorescence is due to the presence of a chromophore intrinsic to the primary structure, thus requiring no additional cofactors. The chromophore is a *p*-hydroxybenzylidene-imidazolinone formed from Ser^{65}-Tyr^{66}-Gly^{67}. Fluorescence is acquired through the creation of the imidazolinone by nucleophilic attack of the amino group of Gly^{67} on the carbonyl group of Ser^{65}, followed by dehydration, and then by oxidation of the hydroxybenzyl side chain of Tyr^{66} by atmospheric oxygen. The β-can crystal structure of wild-type GFP consists of an 11-stranded β-barrel with an α-helix running up the axis of the cylinder. The chromophore is attached to the α-helix and buried in the centre of the cylinder (Ormö et. al. 1996, Tsien 1998). The *in vitro* spectral properties of GFP are

influenced by temperature, ionic strength, protein concentration, and pH (Ward et. al. 1982, Campbell and Choy 2001).

The development of GFP as a reporter provides a unique opportunity to examine the apparently sub-optimal expression of glucocerebrosidase in the *Pichia* system. Since several studies have suggested that organisms do not display a random pattern of synonymous codon usage (Nakamura et. al. 1999) and that disparate patterns of codon bias in the transgene and expression host can have a significant impact on levels of recombinant protein production, two GFP variants were employed as reporters in the *Pichia* system: red-shifted GFP (RSGFP) and enhanced GFP (EGFP). Both RSGFP and EGFP have excitation wavelengths shifted toward the red end of the spectrum, but they exhibit different fluorescence intensities and different codon patterns. RSGFP fluoresces 4-6-fold more brightly than wildtype GFP and displays a codon pattern closer to wildtype GFP and yeast. EGFP fluoresces 35-fold brighter than wildtype GFP and is human codon-optimized. It has been suggested that EGFP is unsuitable for expression in yeast host systems due to preferred codon discrepancies (Clontech 2001). By employing both forms of GFP in creating glucocerebrosidase/GFP chimerae, one can assess whether the more intense fluorescence of EGFP would prove more beneficial in tracking chimeric glucocerebrosidase expression, or whether the greater predicted expression of RSGFP would provide greater production of (and more insight into) glucocerebrosidase. In this manner, one can determine which step (transcription, translation, or purification) provides the greatest obstacle to eventual production and isolation of glucocerebrosidase in the *Pichia* system.

2.2 Methods and Materials

2.2.1 Escherichia coli and Pichia pastoris Strains

The *Escherichia coli* cell line *TOP10F'* used to create the recombinant vectors was obtained from Invitrogen (Carlsbad, CA). The *Pichia pastoris GS115* cell line used for heterologous protein expression was also from Invitrogen (Carlsbad, CA). This strain permits both gene replacement of the yeast *AOX1* gene with one's gene of interest or gene insertion flanking the yeast *AOX1* gene which leaves the host *AOX1* gene intact. The former results in a phenotype of Mut^s (methanol utilization slow) in which cells exhibit slow growth on methanol due to the loss of the *AOX1* gene. The latter results in a Mut^+ phenotype (methanol utilization wildtype) in which rapid growth occurs in the presence of methanol since the *AOX1* gene remains intact and functional. The plasmids in this dissertation exhibit the Mut^+ phenotype.

2.2.2 Plasmid Construction

All expression constructs utilized the pPICZαA vector (Invitrogen, Carlsbad, CA) which contains the *Saccharomyces cerevisae* α-factor secretion signal to direct protein passage out of cells. All primer information is contained in Table 2.1. Plasmid schemas are contained in Figure 2.1. All PCR-amplifications utilized *Pfu* enzyme. To create the recombinant plasmid containing the RSGFP insert, the RSGFP cDNA from pRSGFP-C1 (donation from Dr. D. Levin) was PCR-amplified by primers A and B which contained *Eco* RI restriction cut sites. Subsequently, this 800 base pair (bp) fragment was digested, purified, and cloned in-frame with the pPICZαA α-factor secretion signal to create pPICZαA·RSGFP.

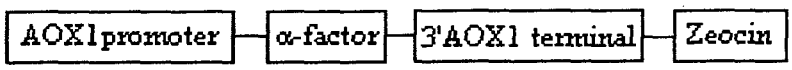
Table 2.1 Primers used in the construction and screening of glucocerebrosidase and green fluorescent protein expression plasmids.

Primer	DNA	Sequence (5' to 3')	Location in cDNA[a]	Orientation
A	RSGFP	TAGAATTCCCGGTCGCCACC ATG	593-615	sense
B	RSGFP	TGTTACAGGGCCCGCGGTTC AGTCGAC	1392-1366	antisense
C	RSGFP	TTTATTGGCCGAGCGGGCCA CCGGTC	581-606	sense
D	EGFP	AACGGTCGAATTCATGGTGA GCAAGGG	665-692	sense
E	EGFP	TATGATCTGAATTGCCGGCC GCTTTACTT	1421-1393	antisense
F	EGFP	TACCTGTGGCATCGCCAGTG GAGTGTGAGCAAGGGCGA	GBA: 1591-1611 EGFP: 679-695[b]	sense
G	GBA	ACTCGAATTCTTCATCTAAG GACCCTGAGG	minus 32-minus 3	sense
H	GBA	ATTTAGGGCCTGCTCGGCCA CTGGCGAT	1628-1601	antisense
I	GBA	TCGCCCTTGCTCACACTCCAC TGGCGATGCCACAGGTA	GBA: 1611-1591 EGFP: 695-679[b]	antisense
J	pPICZαA	GACTGGTTCCAATTGACAAG C	855-875	sense
K	pPICZαA	GCAAATGGCATTCTGACATC C	1443-1423	antisense

[a] Numbers for RSGFP and EGFP correspond to those in the Clontech manuals (Palo Alto, CA). For GBA, the first base of the upstream initiator codon is #1 (Sorge et. al. 1987). For pPICZαA, numbers correspond to those in the pPICZα Manual (Invitrogen, Carlsbad, CA).

[b] Linker primers contained both GBA and EGFP cDNA for the purposes of linking the two cDNAs together.

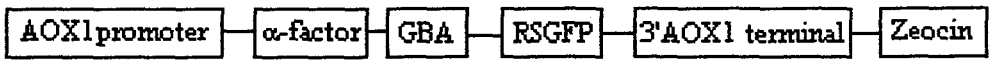
Note. RSGFP= red-shifted green fluorescent protein; EGFP= enhanced green fluorescent protein; GBA= glucocerebrosidase; pPICZαA= pPICZαA expression plasmid.



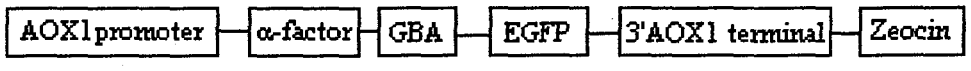
(a) pPICZα A

(b) pPICZα A•RSGFP

(c) pPICZα A•EGFP



(d) pPICZα A•GBA•RSGFP



(e) pPICZα A•GBA•EGFP

Figure 2.1. *Pichia pastoris* plasmid schemas. Abbreviations are as follows: AOX1 promoter (alcohol oxidase 1 promoter), α-factor (α -factor secretion signal), 3'AOX1 terminal (3' terminal of alcohol oxidase 1), Zeocin (selectable marker for Zeocin™ antibiotic), RSGFP (red-shifted green fluorescent protein), EGFP (enhanced green fluorescent protein), GBA (glucocerebrosidase).

To create the recombinant plasmid containing the 730 bp EGFP insert, the EGFP cDNA from pEGFP-N1 (Clontech, Palo Alto, CA) was PCR-amplified by primers D and E which contained *Eco* RI restriction sites. Creation of the desired pPICZαA·EGFP plasmid was then performed in the same manner as pPICZαA·RSGFP.

To create the double insert plasmid containing both glucocerebrosidase and RSGFP, GBA cDNA was first amplified with primers G and H which contained an *Eco* RI restriction site and an *Sfi* I site, respectively, resulting in a 1.6 kb insert. The stop codon of GBA was mutated in primer H to permit continuous read-through to RSGFP. RSGFP was amplified with primers C and B, which contained an *Sfi* I site and an *Eco* RI site, respectively. Both inserts were digested with *Sfi* I, ligated together, and amplified unpurified together as one continuous insert using primers G and B. Since more than one band resulted from this crude PCR, the products were electrophoresed in a 1% low-melting point agarose gel and the band corresponding to the ligated inserts was excised. This band was melted at 75°C in a waterbath and 5 µl were used as template for a core PCR to amplify the GBA·RSGFP insert. This insert was digested with *Eco* RI, purified, and cloned in-frame with the pPICZαA α-factor secretion signal to create pPICZαA·GBA·RSGFP.

To engineer the double insert plasmid containing both glucocerebrosidase and EGFP, GBA was first amplified with primers G and I, creating an insert with no stop codon and a small 3' cDNA segment complementary to the 5' end of EGFP. EGFP was amplified with primers F and E, creating an insert with a small 5' cDNA segment complementary to the 3' end of GBA. Both inserts were then fused together and amplified with primers G and E by overlap extension adapted from Horton et. al. (1990).

The GBA·EGFP insert was then digested with *Eco* RI, purified, and cloned in-frame into pPICZαA.

2.2.3 *Bacterial Transformation*

To make cells competent, 50 µl of frozen *Escherichia coli TOP10F'* cells were used to inoculate 1.5 ml of low salt LB (LSLB; 1% tryptone, 0.5% yeast extract, 0.5% NaCl)) and grown in a 37°C shaking waterbath overnight. This culture was then used to inoculate 200 ml LSLB and grown to an OD_{600} of 0.55. The cells were cooled on ice for 10 minutes before being centrifuged for 20 minutes at 4 °C (3000 x *g*). The cells were washed twice with 200 ml ice-cold sterile deionized water. A volume of sterile deionized water equal to that of the pellet volume was added prior to pellet resuspension.

Once the cells were competent, 45 µl were added to 1 µl of each ligation, gently mixed, transferred to a 0.1 cm cuvette, and pulsed at 1.5 kV. One (1) ml of SOC medium (2% tryptone, 0.5% yeast extract, 10 *mM* NaCl, 2.5 *mM* KCl, 10 *mM* $MgCl_2 \cdot 6H_20$, 10 *mM* $MgSO_4 \cdot 7H_20$) was added prior to a 30-60 minute incubation at 37 °C.

Transformed cells were spread on LSLB plates containing 25 µg/ml Zeocin™ and incubated at 37 °C overnight. Apparent positive colonies were masterplated on LSLB plates containing 25 µg/ml Zeocin™ and again incubated overnight at 37 °C. Colonies demonstrating growth on the masterplates were screened according to the protocol of Campbell and Choy (2000) using primers J and K for pPICZαA, primers A and B for pPICZαA·RSGFP, primers D and E for pPICZαA·EGFP, primers G and B for pPICZαA·GBA·RSGFP, and primers G and E for pPICZαA·GBA·EGFP. A minimum of

two true positive clones per desired vector was sequenced (Koop DNA Sequencing Service Laboratory, Victoria, BC) to confirm absence of mutations.

A loop of each sequenced true positive was used to inoculate 50 ml of LSLB supplemented with 25 µg/ml Zeocin™ and subjected to shaking overnight at 37 °C. The plasmids were isolated by a phenol:chloroform method modified from Sambrook and Russell (2001). Briefly, 50 ml of cells were pelleted, resuspended with 4 ml Cell Resuspension Solution (50 *mM* Tris-HCl, pH 7.5, 10 *mM* EDTA, 100 µg/ml RNase A), and lysed with 4 ml Cell Lysis Solution (0.2 *M* NaOH, 1% SDS). Four (4) ml Cell Neutralization Solution (1.32 *M* potassium acetate, pH 4.8) was added and the mixture was centrifuged for 15 minutes at 10,000 x *g* before being treated with RNase A for 1 hour. All subsequent centrifugation steps were at 10,000 x *g* and 4 °C. The supernatant was mixed for 5 minutes with an equal volume of buffered phenol and centrifuged for 5 minutes. The resulting supernatant was mixed for 5 minutes with an equal volume of 24:1 chloroform:isoamyl alcohol and centrifuged for 5 minutes. Ethanol precipitation of the plasmids within the final supernatant was performed according to Sambrook and Russell (2001). Final plasmid concentrations were determined via spectrophotometry.

2.2.4 Pichia pastoris Transformation

Ten (10) µg of plasmid DNA were linearized with *Bst* XI and ethanol precipitated according to Sambrook and Russell (2001) prior to electroporation into *Pichia pastoris GS115* cells as described in the Invitrogen pPICZα Manual (Carlsbad, CA). Recombinant clones were selected on YPDS plates (1% yeast extract, 2% peptone, 1*M* sorbitol, 2% agar) containing 25 µg/ml Zeocin™, masterplated, and grown for 4 days at 30 °C.

Apparent positive clones were screened by the direct yeast PCR method using primers J and K as noted in the Invitrogen *Pichia* Expression Manual (Carlsbad, CA).

2.2.5 Pichia pastoris Induction and Heterologous Protein Expression

For expression studies, the procedure was modified from the Invitrogen *Pichia* Expression Manual (Carlsbad, CA). A single colony from the masterplate of each desired recombinant vector was used to inoculate 25 ml buffered glycerol-complex medium (BMGY; 0.1*M* sodium citrate, pH 5.5, 2% peptone, 1% yeast extract, 1.34% yeast nitrogen base, 1% glycerol, 4 x 10^{-5}% biotin), which was then grown with shaking (280 rpm) at 30 °C until the culture reached an OD_{600} of 2-6. The cells were resuspended in buffered methanol-complex medium (BMMY; 0.1*M* sodium citrate, pH 5.5, 2% peptone, 1% yeast extract, 1.34% yeast nitrogen base, 0.5% methanol, 4 x 10^{-5}% biotin) to a final OD_{600} of 1.0. The culture was transferred to a 500 ml baffled flask and grown at 30 °C (280 rpm) for 72 hours with methanol added to 0.5% and DTT added to 10 *mM* every 24 hours.

2.2.6 Protein Analysis

Every 12 hours, a 15 µl aliquot of each expression culture were removed and analyzed by fluorescence microscopy. Presence or absence of green fluorescence was recorded and photographed with both a Zeiss Epifluorescence microscope and a Zeiss LSM410 confocal microscope.

Additionally, 5 ml samples were taken every 12 hours. The sample cells were pelleted for 3 minutes at 10,000 x *g* prior to flash-freezing of the supernatant and

subsequent storage at –20 °C. For sodium dodecyl-sulfate polyacrylamide gel electrophoresis (SDS-PAGE), 20 µl of sample supernatant were thawed, removed, mixed with 10 µl of 3x SDS loading buffer, boiled 5 minutes, and centrifuged 5 minutes at 10,000 x *g*. This was loaded, along with controls and a protein standard, in a 10% tris-glycine gel prior to electrophoresis. For silver-staining, the gel was microwaved for 90 seconds in fixative (50% methanol, 12% acetic acid, 0.1% formaldehyde), then for 90 seconds in 50% ethanol. The gel was subsequently pretreated in 0.02% sodium thiosulfate pentahydrate for 90 seconds in the microwave, washed in deionized water for 90 seconds at room temperature, stained (2 mg/ml silver nitrate, 0.075% formaldehyde) by twice microwaving for 40 seconds, then developed (60 mg/ml sodium carbonate, 0.05% formaldehyde, 0.002% sodium thiosulfate pentahydrate) at room temperature for 15-30 minutes. The reaction was stopped in 50% methanol.

For western blotting, proteins were electroblotted from SDS-PAGE gels onto Hybond-P PVDF membrane (Amersham, Piscataway, NJ) overnight at 20 V in 10% methanol transfer buffer (25 *mM* Tris-HCl, 0.2*M* glycine) using a Mini-protean II Electroblot Apparatus (BioRad, Hercules, CA). All following steps were performed at room temperature with gentle agitation. PVDF membranes were washed for 5 minutes in TTBS (20 *mM* Tris-HCl, pH 7.5, 0.05% Tween-20, 500 *mM* sodium chloride) and blocked for 1 hour in TTBS with 7.5% dry skim milk powder. A 1:1000 dilution of anti-GFP antibody pre-conjugated to horse radish peroxidase (GFP-HRP, Clontech, Palo Alto, CA) was added to the blocking solution and incubated 1 hour. Membranes were washed 4 times with TTBS, incubated with ECL+ chemifluorescent reagent (Amersham,

Piscataway, NJ) for 5 minutes, and visualized on a Molecular Dynamics Storm 860 Phosphorimager (Molecular Dynamics, Sunnyvale, CA).

For dot blots examining GFP expression, PVDF membranes were gridded, washed in 100% methanol, and rinsed twice with deionized water. Sample (2.5 µl) was dotted in each square prior to drying of the membrane. The membrane was then submerged in 100% methanol and washed twice with deionized water before following the aforementioned post-transfer western blot protocol. For dot blots examining glucocerebrosidase expression, a 1:400 dilution of a mouse monoclonal glucocerebrosidase-specific primary antibody (donation from Dr. E. Beutler) was incubated for 1 hour post-blocking. The membrane was washed 4 times for 5 minutes with TTBS and subsequently incubated in a 1:6000 dilution of horse radish peroxidase conjugated goat anti-mouse secondary antibody (Clontech, Palo Alto, CA) in 7.5% skim milk powder/TTBS. This was followed by 4 washes (5 minutes each) with TTBS and chemifluorescent detection.

2.2.7 Codon Adaptation Index Calculations for Green Fluorescent Protein Variants

Based on the calculated preferred codon set for *Pichia pastoris* (G. Sinclair, pers. comm.), codon adaptation index (CAI) values were calculated for both RSGFP and EGFP through Codon W (John Peden, Oxford University). For comparison, CAI values for wildtype and yeast-optimized GFP were also calculated. Codon W is an integrated codon bias and correspondence analysis program capable of performing multivariate analyses of codon usage. CAI values are calculated by determining the relative adaptiveness of each codon in a gene in comparison with a preferred codon set. Sequences for each GFP

variant were entered with position #1 corresponding to the "A" of the initiator "ATG" and without spaces.

2.3 Results

2.3.1 Plasmid Construction and Bacterial Transformation

The following plasmids were successfully used to transform *TOP10F'* cells: pPICZαA, pPICZαA·RSGFP, pPICZαA·EGFP, pPICZαA·GBA·RSGFP, and pPICZαA·GBA·EGFP. Figure 2.2 shows bands within a 0.7% agarose gel corresponding to true positive clones for each plasmid following amplification during the screening procedure. Expected band sizes are as follows: 0.6 kb for pPICZαA, 0.8 kb for pPICZαA·RSGFP, 0.7 kb for pPICZαA·EGFP, 2.4 kb for pPICZαA·GBA·RSGFP, and 2.3 kb for pPICZαA·GBA·EGFP. Sequencing results confirmed the absence of mutations in all vectors.

2.3.2 Pichia pastoris Transformation

Confirmed plasmid true positives were isolated from 50 ml large-scale cultures and successfully introduced into *Pichia pastoris GS115* cells via electroporation. Figure 2.3 illustrates positive direct yeast PCR screening results. Bands in the 0.7% agarose gel correspond to appropriate cassette sizes for each vector. Expected band sizes are as follows: 0.6 kb for pPICZαA, 1.4 kb for pPICZαA·RSGFP, 1.3 kb for pPICZαA·EGFP, 2.9 kb for pPICZαA·GBA·RSGFP, and 2.9 kb for pPICZαA·GBA·EGFP An additional 2.2 kb band corresponding to the *AOX1* gene of the *GS115* strain is also present.

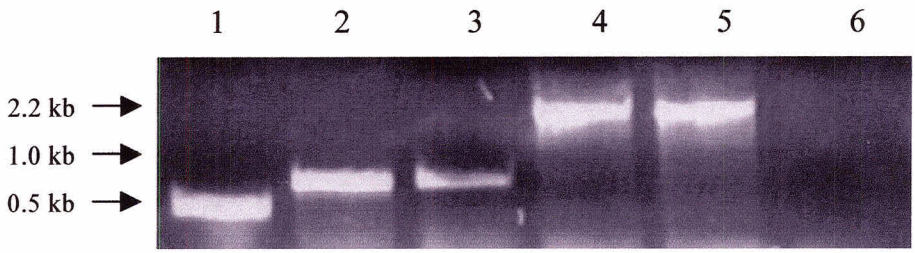


Figure 2.2. Positive clones from transformed *TOP10F'* cells screened by PCR-amplification and subsequent electrophoresis in a 0.7% agarose gel. From left to right (lanes 1-6): pPICZαA, pPICZαA•EGFP, pPICZαA•RSGFP, pPICZαA•GBA•RSGFP, pPICZαA•GBA•EGFP, negative (no DNA) control. Abbreviations are as follows: EGFP (enhanced green fluorescent protein), RSGFP (red-shifted green fluorescent protein), GBA (glucocerebrosidase).

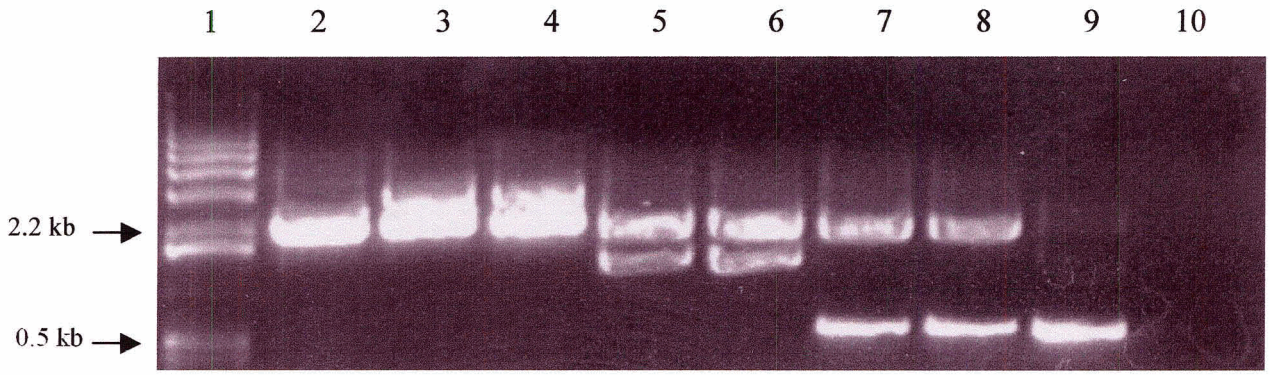


Figure 2.3. *Pichia pastoris* direct yeast PCR products electrophoresed in a 0.7% agarose gel. From left to right (lanes 1-10): lambda kb ladder, *GS115* strain (no plasmid), pPICZαA•GBA•RSGFP, pPICZαA•GBA•EGFP, pPICZαA•RSGFP, pPICZαA•EGFP, pPICZαA (no insert), pPICZαA (no insert), pPICZαA (DNA not in yeast), and negative (no DNA) control. Abbreviations are as follows: GBA (glucocerebrosidase), RSGFP (red-shifted green fluorescent protein), EGFP (enhanced green fluorescent protein). The consistent 2.2 kb band in the lanes 2-8 represents the alcohol oxidase 1 (*AOX1*) gene naturally present in the *GS115* strain.

2.3.3 Pichia pastoris Heterologous Protein Expression and Protein Analysis

Green fluorescence was observed in both the pPICZαA·RSGFP and pPICZαA·EGFP cultures at each successive 12 hour time point starting at 24 hours post-induction. No fluorescence was observed in either the pPICZαA·GBA·RSGFP or the pPICZαA·GBA·EGFP cultures at any time point. As expected, no fluorescence was observed in the vector-only pPICZαA and cells-only *GS115* strain controls (Figure 2.4a, Figure 2.4b).

SDS-PAGE followed by silver staining revealed a ladder of proteins present in the induction medium for all cultures. A 27,000 MW band representing EGFP was present in the pPICZαA·EGFP-expressing culture. No other clear bands corresponding to the expected protein molecular weights were visible (Figure 2.5).

Protein blot analysis indicated cross-reactivity with both expressed EGFP and RSGFP proteins at approximately 27,000 MW. Greater intensity of cross-reactivity was observed with EGFP, corroborating the SDS-PAGE silver stain results. No bands of the expected 89,000 MW were visible for pPICZαA·GBA·EGFP or pPICZαA·GBA·RSGFP. The negative control cultures, pPICZαA and *GS115*, did not demonstrate any cross-reactivity (Figure 2.5).

Protein dot blots utilizing the GFP antibody demonstrated cross-reactivity with EGFP- and RSGFP-expressing cultures at both 48 hours and 72 hours post-induction. No dots were visible for the chimeric GBA-expressing cultures or negative controls (Figure 2.6). Protein dot blots utilizing the GBA antibody showed no cross-reactivity with any of the expression cultures (Figure 2.7). Taken together, these results indicate successful EGFP and RSGFP expression, but no detectable GBA·RSGFP or GBA·EGFP production.

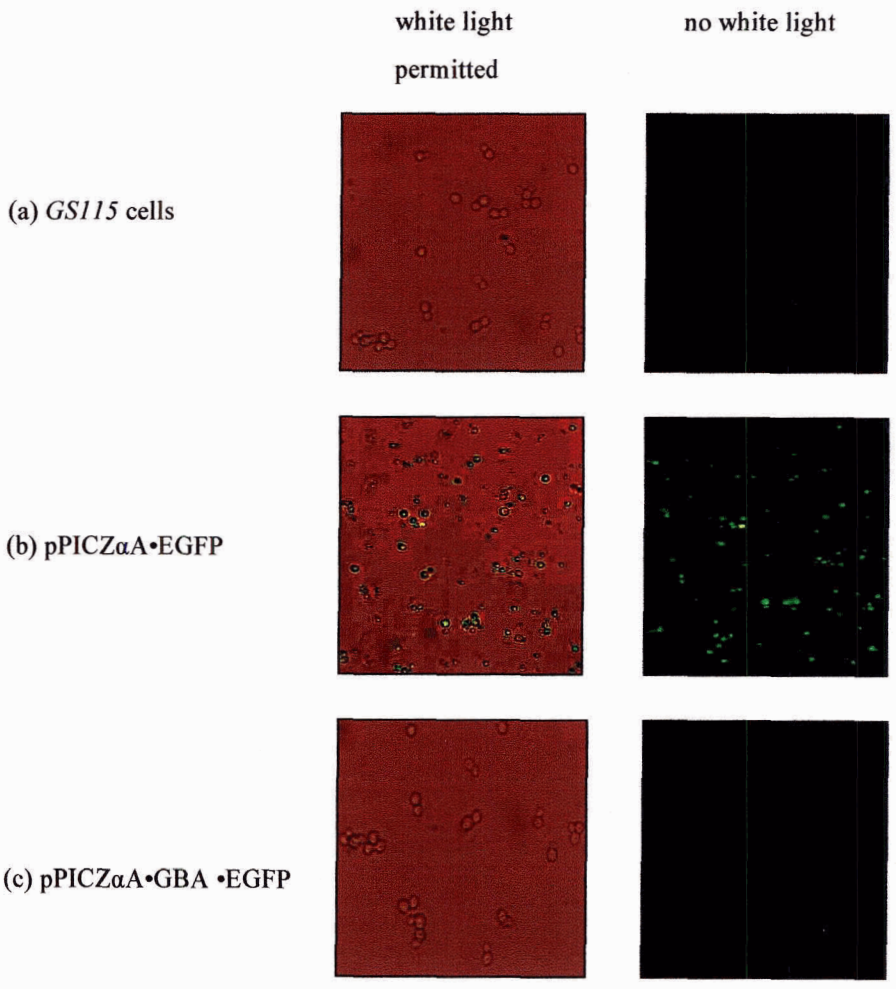


Figure 2.4a. Heterologous protein expression at 72 hours post-induction in *Pichia pastoris.* Abbreviations are as follows: EGFP (enhanced green fluorescent protein), GBA (glucocerebrosidase).

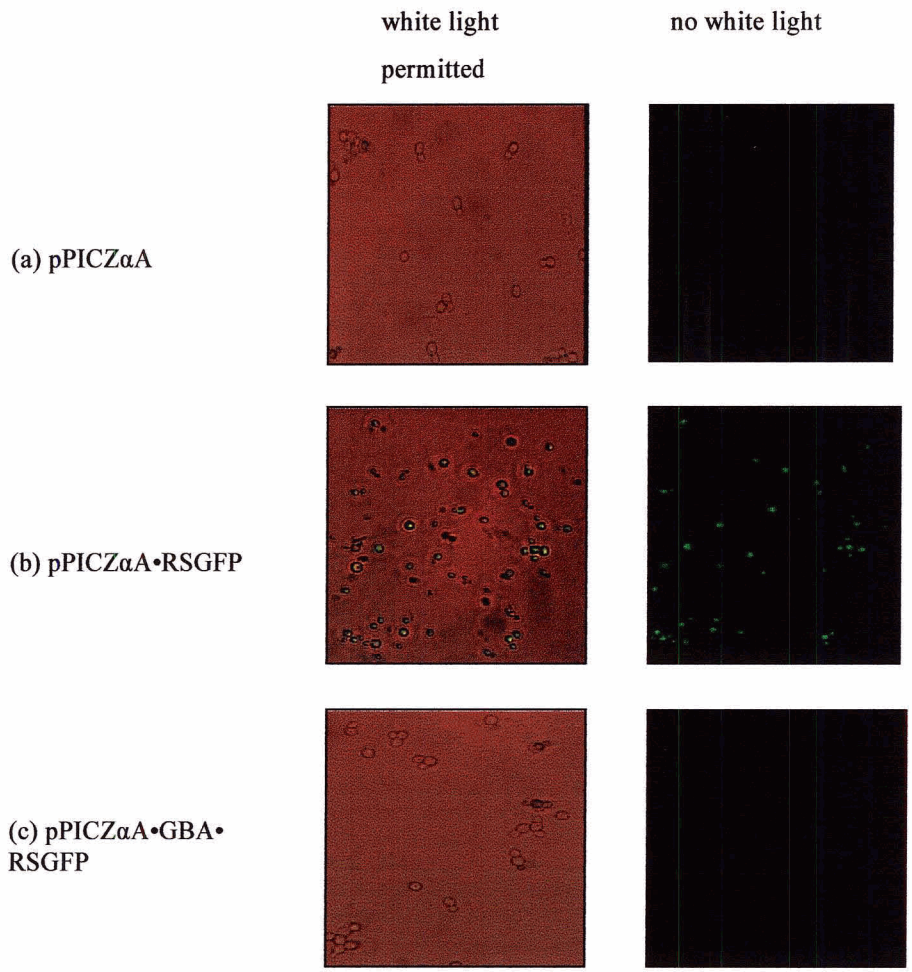


Figure 2.4b. Heterologous protein expression at 72 hours post-induction in *Pichia pastoris.* Abbreviations are as follows: RSGFP (red-shifted green fluorescent protein), GBA (glucocerebrosidase).

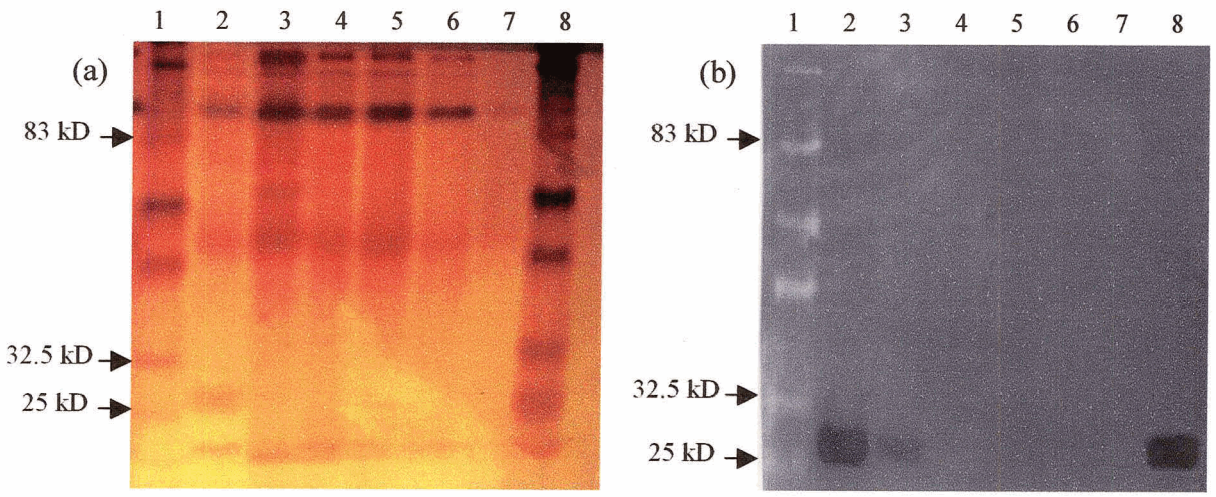


Figure 2.5. Silver-stained SDS-PAGE gel (a) and replica western blot (b) of secreted proteins in medium from induced *Pichia pastoris* cultures 72 hours post-induction. Immunodetection was performed using an anti-GFP antibody. (a) From left to right (lanes 1-8): pre-stained protein marker, pPICZαA•EGFP, pPICZαA•RSGFP, pPICZαA•GBA•EGFP, pPICZαA•GBA•RSGFP, pPICZαA, *GS115* cells, pre-stained protein marker. (b) From left to right (lanes 1-8): pre-stained protein marker, pPICZαA•EGFP, pPICZαA•RSGFP, pPICZαA•GBA•EGFP, pPICZαA•GBA•RSGFP, pPICZαA, *GS115* cells, and GFP positive control. Abbreviations are as follows: EGFP (enhanced green fluorescent protein), RSGFP (red-shifted green fluorescent protein), GBA (glucocerebrosidase), kD (kilodalton, equal to 1000 MW).

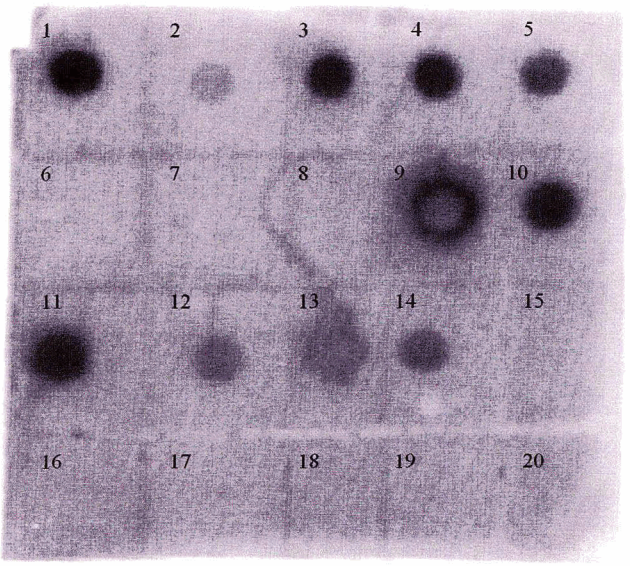


Figure 2.6. Protein dot blot of secreted medium from induced *Pichia pastoris* cultures. Immunodetection was performed using an anti-GFP antibody. Boxes 1-20 are as follows: (1) positive GFP control (2) pPICZαA•EGFP *(24h)* (3) pPICZαA•EGFP *(48h)* (4) pPICZαA•RSGFP *(24h)* (5) pPICZαA•RSGFP *(48h)* (6) pPICZαA *(48h)* (7) pPICZαA•GBA•EGFP *(24h)* (8) pPICZαA•GBA•EGFP *(48h)* (9) pPICZαA•EGFP *(24h)* (10) pPICZαA•EGFP *(48h)* (11) pPICZαA•EGFP *(72h)* (12) pPICZαA•RSGFP *(24h)* (13) pPICZαA•RSGFP *(48h)* (14) pPICZαA•RSGFP *(72h)* (15) pPICZαA•GBA•EGFP *(24h)* (16) pPICZαA•GBA•EGFP *(48h)* (17) pPICZαA•GBA•EGFP *(72h)* (18) pPICZαA•GBA•RSGFP *(24h)* (19) pPICZαA•GBA•RSGFP *(48h)* (20) pPICZαA•GBA•RSGFP *(72h)* . Abbreviations are as follows: 24h (24 hours post-induction), 48h (48 hours post-induction), 72h (72 hours post-induction), EGFP (enhanced green fluorescent protein), RSGFP (red-shifted green fluorescent protein), GBA (glucocerebrosidase).

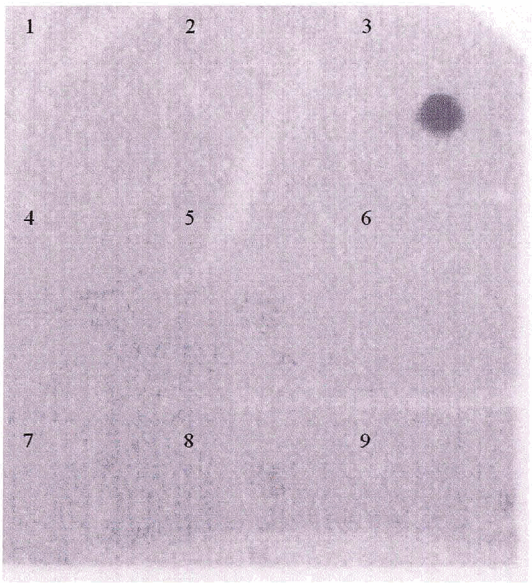


Figure 2.7. Protein dot blot of secreted medium from induced *Pichia pastoris* cultures. All samples were taken at 72 hours post-induction unless otherwise indicated. Immunodetection was performed using a human GBA-specific antibody. Boxes 1-9 are as follows: (1) pPICZαA•EGFP (2) pPICZαA (3) positive control (GBA purified from *Sf9* insect cell medium) (4) pPICZαA•RSGFP (5) pPICZαA•GBA•EGFP 48 hours post-induction (6) pPICZαA•GBA•RSGFP 48 hours post-induction (7) negative control (*GS115* cells) (8) pPICZαA•GBA•EGFP (9) pPICZαA•GBA•RSGFP. Abbreviations are as follows: EGFP (enhanced green fluorescent protein), RSGFP (red-shifted green fluorescent protein), GBA (glucocerebrosidase).

2.3.4 Green Fluorescent Protein Codon Adaptation Index Calculations

Highly adapted genes with codon choice identical to that preferred by the host organism yield CAI values of 1, while genes with low overall magnitudes of bias or bias toward a different set of codons yield values approaching zero. RSGFP had a CAI value of 0.423, while EGFP had a CAI value of 0.380. CAI values for wildtype GFP and yeast-optimized GFP were also calculated, yielding values of 0.439 and 0.639, respectively.

2.4 Discussion

Due to the exorbitant cost of enzyme replacement therapy for Gaucher patients, the Choy lab began testing the *Pichia pastoris* system as a possible cost-effective method of producing glucocerebrosidase. As a eukaryote, *P. pastoris* provides protein processing, folding, and posttranslational modification, while being faster, easier, and less expensive to use than mammalian expression systems (Higgins and Cregg 1998). The presence of a tightly controlled alcohol oxidase 1 gene also permits inducible, enhanced expression of a gene of interest, representing up to 30% of total soluble protein in fermenter cultures (Lin Cereghino and Cregg 2000). Currently, more than 400 proteins have been expressed in the *P. pastoris* system (Lin Cereghino et. al. 2002). However, despite the successful expression of numerous proteins by other laboratories, the Choy lab was unable to produce significant amounts of glucocerebrosidase in this system. Alterations in growth conditions, plasmid choice, and cassette copy numbers failed to result in desired levels of expression. Therefore, in order to provide insight into optimizing glucocerebrosidase production, green fluorescent protein was fused to glucocerebrosidase, serving as a visible *in vivo* fluorescent marker.

Results indicated that glucocerebrosidase chimerae were not produced at detectable levels. No fluorescence was observed in either the glucocerebrosidase/red-shifted green fluorescent culture or the glucocerebrosidase/enhanced green fluorescent protein culture. No bands corresponding to the expected 89,000 MW were visible in SDS-PAGE gels. Moreover, no cross-reactivity was observed in western blots or protein dot blots. Thus, the results of this study corroborated the difficulty of glucocerebrosidase production previously noted in the Choy lab. Since no fluorescence was observed *in vivo* prior to protein isolation from *Pichia* medium, the problems associated with glucocerebrosidase production occur before the protein isolation and purification procedures. In other words, it is not a low yield recovery post-isolation that is causing the low enzyme levels, but rather a hindrance at the transcriptional or translational levels. This study, therefore, suggests that the *P. pastoris* system is currently not an appropriate choice for glucocerebrosidase production for the purpose of enzyme replacement therapy. Experiments in Chapters 3 and 4 of this dissertation further examine the issue of glucocerebrosidase biosynthesis.

In contrast, both variants of green fluorescent protein were successfully produced in the *Pichia* system. Both RSGFP and EGFP exhibited fluorescence at each time point after 24 hours post-induction. Though only a clear band representing EGFP appeared in the silver-stained SDS-PAGE gel, both EGFP and RSGFP showed appropriate cross-reactivity in the western blots and protein dot blots. Unexpectedly, EGFP appeared to be secreted in a greater quantity than RSGFP. CAI values for RSGFP and EGFP are 0.423 and 0.380, respectively, suggesting that RSGFP should be produced at higher levels than EGFP. The results of this study, however, do not support previous suggestions that the

expression of EGFP should be avoided in yeast (Clontech, Palo Alto, CA), nor do they corroborate the codon bias theory (Campbell and Choy 2002).

A number of possible reasons exist to explain this discrepancy. Kurland (1991) notes that "a heterologous gene is not necessarily expressed at a low level simply because it is made up of codons that are infrequently translated by the host cell." The codon sequences within a given genome may be locally biased. Thus, discontinuities of base composition between local domains may lead to a mosaic arrangement of codon preferences (Kurland 1991). Likewise, classes of genes within the same genome that are physiologically regulated to different expression levels may have class-specific codon preferences (Gouy and Gautier 1982). Furthermore, intragenic codon biases exist. In *E. coli* and *Saccharomyces cerevisae*, there is a codon bias in the initial sequences of genes which for major proteins is strikingly different from their downstream bias (Bulmer 1988, Kurland 1991). Moreover, RNA or protein products expressed from heterologous genes may be unstable in the host cell (Kurland 1991). Thus, these factors may play a role in the disparate expression of EGFP and RSGFP observed in the *Pichia pastoris* system.

2.5 References

Allen, G.J., Kwak, J.M., Chu, S.P., Llopis, J., Tsien, R.Y., Harper, J.F., and Schroeder, J.I. (1999). Cameleon calcium indicator reports cytoplasmic calcium dynamics in *Arabidopsis* guard cells. *Plant J.* 19: 735-747.

Beutler, E. and Grabowski, GA. (2001). Gaucher disease. In Scriver, C.R., Beaudet, A.L, Sly, W.S., and Valle, D (eds.), *The Metabolic and Molecular Bases of Inherited Disease* (vol. 3, 8th ed., pp. 3635-3668). New York: McGraw-Hill.

Bulmer, M. (1988). Codon usage and intragenic position *J. Theor. Biol.* 133: 67-71.

Campbell, T.N. and Choy, F.Y.M. (2000). Large-scale colony screening and insert orientation determination using PCR. *BioTechniques* 30: 32-34.

Campbell, T.N. and Choy, F.Y.M. (2001). The effect of pH on green fluorescent protein: a brief review. *Molecular Biology Today*: 2: 1-4.

Campbell, T.N. and Choy, F.Y.M. (2002). Expression of two green fluorescent protein variants in citrate-buffered media in *Pichia pastoris. Anal. Biochem.* 311: 193-195.

Cha, H.J., Dalal, N.G., Vakharia, V.N., and Bentley, W.E. (1999). Expression and purification of human interleukin-2 simplified as a fusion with green fluorescent protein in suspended Sf-9 insect cells. *J. Biotechnol.* 69: 9-17.

Chalfie, M., Tu, Y., Euskirchen, G., Ward, W.W., and Prasher, D.C. (1994). Green fluorescent protein as a marker for gene expression. *Science* 263: 802-805.

Clontech. (2001). *Living Colors® User Manual (PT2040-1).* Palo Alto: Clontech.

Dąbrowski, S., Brillowska, A., and Kur, J. (1999). Use of the green fluorescent protein variant (YFP) to monitor MetArg human proinsulin production in *Escherichia coli. Prot. Ex. Pur.* 16: 315-323.

Dhandayuthapani, S., Via, L.E., Thomas, C.A., Horowitz, P.M., Deretic, D., and Deretic, V. (1995). Green fluorescent protein as a marker for gene expression and cell biology of mycobacterial interactions with macrophages. *Mol. Microbiol.* 17: 901-912.

Genzyme Corporation. (2003). About Cerezyme: enzyme replacement therapy [WWW document]. URL http://www.cerezyme.com/healthcare/about/p_hc_aboutcz.asp

Gerdes, H-H., and Kaether, C. (1996). Green fluorescent protein: applications in cell biology. *FEBS Lett.* 389: 44-47.

Gouy, M. and Gautier, C. (1982). Codon usage in bacteria: correlation with gene sensitivity. *Nucleic Acids Res.* 10: 7055-7074.

Grabowski, G.A., Leslie, N., and Wenstrup, R. (1998). Enzyme therapy for Gaucher disease: the first 5 years. *Blood Rev.* 12: 115-133.

Harada, M., Sakisaka, S., Terada, K., Kimura, R., Kawaguchi, T., Koga, H., Taniguchi, E., Sasatomi, K., Miura, N., Suganuma, T., Fujita, H., Furuta, K., Tanikawa, K., Sugiyama, T., and Sata, M. (2000). Role of AATP7B in biliary copper excretion in a human hepatoma cell line and normal rat hepatocytes. *Gastroenterology* 118: 921-928.

Higgins, D.R. and Cregg, J.M. (1998). *Methods in Molecular Biology: Pichia Protocols.* Totowa: Humana Press.

Horton, R.M., Cai, Z., Ho, S.N., and Pease, L.R. (1990). Gene splicing by overlap extension: tailor-made genes using the polymerase chain reaction. *BioTechniques* 8: 528-535.

Kneen, M., Farinas, J., Li, Y., and Verkman, A.S. (1998). Green fluorescent protein as a noninvasive intracellular pH indicator. *Biophys. J.* 74: 1591-1599.

Kurland, C.G. (1991). Codon bias and gene expression. *FEBS Lett.* 285: 165-169.

Li, L., Li, Y., Lim, T.M., and Pan, S.Q. (1999). GFP-aided confocal laser scanning microscopy can monitor *Agrobacterium tumefaciens* cell morphology and gene expression associated with infection. *FEMS Microbiol. Lett.* 179: 141-146.

Lin Cereghino, G.P., Lin Cereghino, J., Ilgen, C., and Cregg, J.M. (2002). Production of recombinant proteins in fermenter cultures of the yeast *Pichia pastoris. Curr. Opin. Biotechnol.* 13: 329-332.

Lin Cereghino, J. and Cregg, J.M. (2000). Heterologous protein expression in the methylotrophic yeast *Pichia pastoris. FEMS Microbiol. Rev.* 24: 45-66.

Miesenböck, G., De Angelis, D.A., and Rothman, J.E. (1998). Visualizing secretion and synaptic transmission with pH-sensitive green fluorescent proteins. *Nature* 349: 192-195.

Miyawaki, A., Llopis, J., Heim, R., McCaffery, J.M., Adams, J.A., Ikura, M., and Tsien, R.Y. (1997). Fluorescent indicators for Ca^{2+} based on green fluorescent proteins and calmodulin. *Nature* 388: 882-887.

Morise, H., Shimomura, O., Johnson, F.J., and Winant, J. (1974). Intermolecular energy transfer in the bioluminescent system of *Aequorea. Biochemistry* 13: 2656-2662.

Nakamura, Y. Gojobori, T., and Ikemura, T. (1999). Codon usage tabulated from the international DNA sequence database; its status 1999. *Nucleic Acids Res.* **27**, 292.

Ormö, M., Cubitt, A.B., Kallio, K., Gross, L.A., Tsien, R.Y., and Remington, S.J. (1996). Crystal structure of the *Aequorea victoria* green fluorescent protein. *Science* 273: 1392-1395.

Palm, G.J. and Wlodawer, A. (1999). Spectral variants of green fluorescent protein. *Methods Enzymol.* 302: 378-394.

Peden, J.F. (1999). *Analysis of codon usage.* Unpublished doctoral dissertation, University of Nottingham, University Park, Nottingham, United Kingdom.

Prasher, D.C., Eckenrode, V.K., Ward, W.W., Prendergast, F.G., and Cormier, M.J. (1992). Primary structure of the *Aequorea victoria* green-fluorescent protein. *Gene* 111: 229-233.

Rizzuto, R., Brini, M., De Gorgi, F., Rossi, R., Heim, R., Tsien, R.Y., and Pozzan, T. (1996). Double labelling of subcellular structures with organelle-targeted GFP mutants *in vivo*. *Curr. Biol.* 6: 183-188.

Rizzuto, R., Brini, M., Pizzo, P., Murgia, M., Pozzan, T. (1995). Chimeric green fluorescent protein as a tool for visualizing subcellular organelles in living cells. *Curr. Biol.* 5: 635-642.

Sambrook, J. and Russell, D.W. (2001). *Molecular Cloning: a Laboratory Manual.* New York: Cold Spring Harbor.

Shimomura, O., Johnson, F.H., and Saiga, Y. (1962). Extraction, purification and properties of Aequorin, a bioluminescent protein from the luminous hydromedusan, *Aequorea*. *J. Cell Comp. Physiol.* 59: 223-239.

Sorge, J.A., West, C., Kuhl, W., Treger, L., and Beutler, E. (1987). The human glucocerebrosidase gene has two functional ATG initiator codons. *Am. J. Hum. Genet.* 41: 1016-1024.

Tsien, R.Y. (1998). The green fluorescent protein. *Annu. Rev. Biochem.* 67: 509-544.

Valdivia, R.H. and Falkow, S. (1996). Bacterial genetics by flow cytometry: rapid isolation of *Salmonella typhimurium* acid-inducible promoters by differential fluorescence induction. *Mol. Microbiol.* 22: 367-378.

Wang, S. and Hazelrigg, T. (1994). Implications for *bcd* mRNA localization from spatial distribution of *exu* protein in *Drosophila* oogenesis. *Nature* 369: 400-403.

Ward, W.W., Cody, C.W., Hart, R.C., and Cormier, M.J. (1980). Spectrophotometric identity of the energy transfer chromophores in *Renilla* and *Aequorea* green-fluorescent proteins. *Photochem. Photobiol.* 31: 611-615.

Ward, W.W., Prentice, H.J., Roth, A.F., Cody, C.W., and Reeves, S.C. (1982). Spectral perturbations of the *Aequorea* green-fluorescent protein. *Photochem. Photobiol.* 35: 803-808.

Wheeler, G.N., Hamilton, F.S., and Hoppler, S. (2000). Inducible gene expression in transgenic *Xenopus* embryos. *Curr. Biol.* 10: 849-852.

Chapter 3 – Examination of Glucocerebrosidase and Enhanced Green Fluorescent Protein Biosynthesis in HeLa Cells

3.1 Introduction

Enzyme replacement therapy, first proposed for Gaucher disease in 1966 (Brady 1966), has become the treatment of choice for Gaucher patients (de Fost et. al. 2003). The expense of production and exorbitant price of treatment, however, has stimulated a search for more efficient and cost-effective methods of production. This has lead to a closer examination of glucocerebrosidase biosynthesis and trafficking. Overexpression of glucocerebrosidase has been attempted in various cell types, ranging from mammalian cells to transgenic tobacco plants (Cramer et. el. 1996, Pasmanik-Chor et. al. 1996, Liu et. al. 1998). To date, however, many unanswered questions still remain. For example, the cause of the low yields of glucocerebrosidase produced in some heterologous protein expression systems has not been determined. As noted in Chapter 2 of this dissertation, the *P. pastoris* system did not produce significant amounts of enzyme.

Another unanswered question is how glucocerebrosidase is trafficked to the lysosome. As stated in Chapter 1, most lysosomal hydrolases utilize a mannose-6-phosphate targeting system for transport to the lysosome. In this system, a signal patch on the hydrolase is recognized by a phosphotransferase which adds an *N*-acetylglucosamine (GlcNAc) -phosphate to one or two mannose residues on each oligosaccharide chain. A phosphoglycosidase then trims off the GlcNAc, creating the mannose-6-phosphate marker while in the *cis* Golgi (Alberts et. al. 1994). Upon entrance into the *trans* Golgi network, these tagged hydrolases are recognized by a mannose-6-phosphate receptor. The

enzyme/receptor complex is shuttled to endosomes via clathrin-coated vesicles for final transport to the lumen of the lysosome (Lemansky et. al. 1989, Glickman and Kornfield 1993). Upon entry into the lysosome, the mannose receptor dissociates and is recycled back to the *trans* Golgi network. Unlike most lysosomal hydrolases, however, glucocerebrosidase does not utilize the mannose-6-phosphate pathway (Beutler and Grabowski 2001).

An alternate mode of lysosomal transport is the cytoplasmic tail targeting system. Acid phosphatase, for example, is rapidly endocytosed from the cell surface in clathrin-coated pits due to a tyrosine-containing internalization signal in its 19 amino acid cytoplasmic tail (Lehmann et. al. 1992). Computer analysis of the glucocerebrosidase core, however, does not indicate any transmembrane domain aside from the signal peptide that is cleaved following transport into the endoplasmic reticulum (Beutler and Grabowski 2001). Accordingly, the mature glucocerebrosidase peptide lacks the appropriate cytoplasmic tail required by such a targeting mechanism.

The role of glycosylation has also been examined with respect to glucocerebrosidase lysosomal transport. The evidence has been mixed: some studies have suggested indirect roles (Aerts et. al. 1986), while others have found no role at all (Leonova and Grabowski 2000). Thus, the mechanism guiding glucocerebrosidase transport to the lysosome remains to be elucidated.

In order to examine glucocerebrosidase production *in vivo* in a higher eukaryotic system and to search for the elusive trafficking mechanism, glucocerebrosidase/green fluorescent protein fusion vectors were created and introduced into HeLa cells. Expression of EGFP-tagged glucocerebrosidase constructs beginning at either the first or

the second ATG initiation site was examined. Transport of full-length glucocerebrosidase chimerae were arrested with temperature and chemical blockages to further examine biosynthetic transport (Figure 3.1).

3.2 Methods and Materials

3.2.1 HeLa Culture

HeLa cells (donation from Dr. T. Pfeiffer) were cultured at 37°C with 5% CO_2 in RPMI 1640 with 10% fetal bovine serum and 5% penicillin/streptomycin (Invitrogen, Carlsbad, CA).

3.2.2 Plasmid Construction

All inserts were cloned into the pEGFP-N1 vector (Clontech, Palo Alto, CA) after digestion with *Eco* RI. Primer information is contained in Table 3.1. All PCR-amplifications utilized *Pfu* enzyme. To create the plasmid containing full-length glucocerebrosidase beginning at the upstream ATG initiation codon (pGBA1·EGFP), glucocerebrosidase template was amplified with primers G and L, digested, purified and cloned in-frame upstream from EGFP. The stop codon of glucocerebrosidase was mutated in primer L to permit continuous read-through to EGFP.

To create the plasmid containing glucocerebrosidase beginning at the second ATG initiation codon (pGBA2·EGFP), glucocerebrosidase template was amplified with primers M and L, digested, purified and cloned in-frame upstream from EGFP.

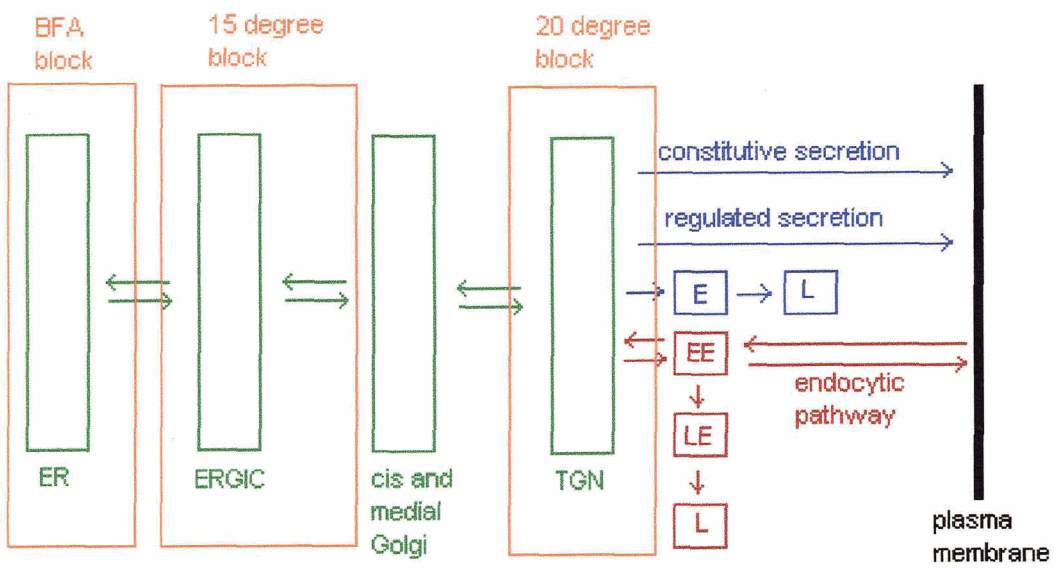


Figure 3.1. Schematic diagram of trafficking among the biosynthetic (green), secretory (blue), and endocytic (red) pathways. Orange rectangles represent locations of blocked traffic under the labelled conditions. Abbreviations are as follows: BFA (brefeldin A), ER (endoplasmic reticulum), ERGIC (endoplasmic reticulum-Golgi intermediate compartment), TGN (trans Golgi network), E (endosome), L (lysosome), EE (early endosome), LE (late endosome). This figure was modified from Campbell and Choy (2001).

Table 3.1. Primers used in the construction and screening of glucocerebrosidase (GBA) and enhanced green fluorescent protein (EGFP) expression plasmids in HeLa cells.

Primer	**DNA**	**Sequence (5' to 3')**	**Location in cDNA[a]**	**Orientation**
G	GBA	ACTCGAATTCTTCATCTAAGGACC CTGAGG	minus 32-minus 3	sense
L	GBA	ATCGAATTCGTCCCTGGCGATGCC AC	1621-1596	antisense
M	GBA	CCTGTGAGGCTGCCAGCCATGAA TTCGGTA	51-80	sense
N	EGFP	ATGGTGAGCAAGGGCGAGGAGCT GT	679-703	sense
P	EGFP	TACAGCTCGTCCATGCCGAGAGT GATCC	1391-1364	antisense

[a] The first base of the GBA upstream initiator codon is designated #1 (Sorge et. al. 1987). Numbers for EGFP correspond to those in Clontech manuals (Palo Alto, CA).

Note. No primer was designated "O" since this may be confused with zero in the text of this dissertation.

3.2.3 Bacterial Transformation

TOP10F' bacterial cells were transformed via electroporation as described in Chapter 2. Once transformed, cells were spread on LB plates containing 30 μg/ml kanamycin and incubated at 37°C overnight. Colonies were screened by PCR-amplification according to the method of Gussow and Clackson (1989) using primers N and P for pEGFP-N1, primers G and P for pGBA1·EGFP, and primers M and P for pGBA2·EGFP. A minimum of two true positive clones per desired vector were sequenced (Koop DNA Sequencing Service Laboratory, Victoria, BC) to confirm absence of mutations.

Large-scale plasmid DNA extractions of sequence-confirmed clones were performed according to the method described in Chapter 2 with one alteration: 30 μg/ml kanamycin was added to culture medium in place of 25 μg/ml Zeocin™ since the pEGFP-N1 plasmid confers kanamycin resistance, not Zeocin™ resistance.

3.2.4 HeLa Transfection

HeLa cells in 60 mm petri dishes were transfected at 80% confluency with FuGENE6 in a 6:2 ratio (μl FuGENE6: μg DNA) according to the manufacturer's instructions (Roche, Indianapolis, IN). For mock-transfected HeLa control cultures, FuGENE6 was added alone. Fetal bovine serum (10%) was added to the dishes 3-5 hours post-transfection. In cultures set aside for microscopic visualization, six sterile 12mm round coverslips were coated with poly-lysine, dried overnight, and placed in the bottom of each 60 mm dish prior to seeding with trypsinized HeLa cells. Cells were grown to 80% confluency and similarly transfected with FuGENE6.

3.2.5 RNA Analysis

Total RNA was extracted 48 hours post-transfection with the RNeasy® Mini Kit according to the manufacturer's instructions for 60mm petri dishes (QIAGEN, Valencia, CA).

For RT-PCR analysis, 1 µl of RNA was used as template for 36 cycles of amplification according to the recommended protocol for Superscript™ One-Step RT-PCR with Platinum®Taq (Invitrogen, Carlsbad, CA). Primers N and P (58°C anneal) were utilized for all reactions. Additionally, primers M and P (58°C anneal) were used in separate reactions for pGBA1·EGFP and pGBA2·EGFP. A control with Platinum® Taq but without reverse transcriptase was included to test whether residual contaminating DNA was present in the isolated RNA samples.

For Northern blots, total RNA concentrations determined by spectrophotometry were standardized. 6 µg of each sample was added to 5X RNA loading dye (0.16% bromophenol blue, 4 *mM* EDTA [pH 8.0], 0.9 *M* formaldehyde, 3.2% formamide, 20% glycerol, 0.4 X MOPS buffer [0.2 *M* 3-{N-morpholino}propanesulfonic acid, 0.08 *M* sodium acetate, 0.01 *M* EDTA; pH 7.0]). This was incubated at 65°C for 5 minutes, cooled on ice for 5 minutes, and loaded in a modified 1% agarose gel (1.0 g agarose, 1X MOPS buffer, 0.2 *M* formaldehyde, 0.1 µg/ml ethidium bromide). Samples were electrophoresed at 70 V for 2 hours in RNA running buffer (1X MOPS buffer, 0.2 *M* formaldehyde). The RNA was transferred to a Hybond-N^+ nylon membrane (Amersham, Piscataway, NJ) by capillary transfer overnight at 4°C in 10X SSC (1.5 *M* sodium chloride, 0.15 *M* sodium citrate). Subsequently, the transferred RNA was cross-linked to the membrane by incubation at 80°C for 2 hours. A DNA probe representing the full-

length EGFP cDNA was PCR-amplified using primers N and P (Table 3.1) and conjugated with alkaline phosphatase according to the AlkPhos Direct Manual (Amersham, Piscataway, NJ). Membranes were hybridized with 100 ng of labelled probe at 55°C overnight and washed as directed. Bands were developed by incubation with ECF chemifluorescent reagent (Amersham, Piscataway, NJ) as directed for 24-48 hours before visualizing with a Molecular Dynamics Storm 860 phosphorimager (Molecular Dynamics, Sunnyvale, CA).

3.2.6 Microscopic Analysis and Antibody Labelling

Every 24 hours, one coverslip per expression culture was removed, mounted on a clean slide with ProLong Antifade (Invitrogen, Carlsbad, CA) and analyzed by fluorescence microscopy. Addition of diluted trypan blue (donation from Dr. R. Burke) in a 1:1 ratio (volume trypan blue: volume cells) was included every second replicate to examine cell viability. Presence or absence of green fluorescence was recorded and photographed with both a Zeiss Epifluorescent microscope and a Zeiss LSM410 confocal microscope.

For antibody double labelling, cells grown on coverslips in 12-well plates were rinsed with 800 µl 1X phosphate-buffered saline (PBS, pH 7.4) 48-72 hours post-transfection. Aside from the primary antibody incubations, all washes and incubations were performed at room temperature. Eight hundred (800) µl of 4% paraformaldehyde was added to the wells, followed by a 5 minute incubation and a rinse with 1X PBS. 800 µl of ice-cold 100% methanol was then added, followed by a 5 minute incubation and two rinses with 1X PBS. The cells were then pre-blocked with 800 µl 5%FBS-PBST (1X

PBS, 0.1% Triton X-100, 5% fetal bovine serum) for 1 hour. Primary antibodies were added according to the dilutions listed in Table 3.2. In each case, the anti-GFP antibody (C) was added along with one of the organelle antibodies (A or B). For negative controls, mock-transfected HeLa cells and an irrelevant primary antibody (E) were used. Primary reactions were incubated at 4°C overnight. Following this incubation, cells were rinsed 3 times for 15 minutes with 1X PBS. Secondary antibodies (D) were then added in dilutions listed in Table 3.2 to 800 µl 5%FBS-PBST and incubated for 2 hours. Cells were then rinsed 3 times with 1X PBS and mounted with ProLong Antifade prior to visualization.

For blockage studies, only cells transfected with pEGFP-N1 and pGBA1·EGFP, as well as mock-transfected HeLa controls, were used. Cells were grown on coverslips in 12-well plates and subjected to 30-90 minutes of one of the following after the addition of 20 *mM* HEPES to each well 48-72 hours post-transfection: incubation at 15°C, incubation at 20°C, addition of 5 µg/ml brefeldin A (BFA) at 37 °C. Following treatment, coverslips were mounted on glass slides with ProLong Antifade and immediately visualized/photographed with either a Zeiss Epifluorescence or Zeiss LSM410 confocal microscope.

3.2.7 Protein Isolation and Analysis

To isolate proteins, medium was removed from HeLa cells in 60 mm dishes 72 hours post-transfection. Cells were washed, trypsinized (0.25% trypsin, 0.03% EDTA), resuspended in medium, and pelleted 10 minutes at 2000 xg. All centrifugation steps were at room temperature. Supernatant was removed and the cells were washed twice

Table 3.2. Antibodies used to visualize glucocerebrosidase and enhanced green fluorescent protein expression in HeLa cells 48-72 hours post-transfection. Abbreviations are as follows: ERGIC (endoplasmic reticulum-Golgi intermediate compartment), ER (endoplasmic reticulum), GFP (green fluorescent protein).

Thesis Label	**Antibody Name**	**Primary or Secondary**	**Source**	**Localization**	**Dilution**	**Donor/ Company of Purchase**
A	anti-giantin	primary	mouse anti-human	Golgi	1:1000	Dr. Hauri
B	anti-ERGIC-53	primary	mouse anti-human	ER and ERGIC	1:1000	Dr. Hauri
C	anti-GFP	primary pre-conjugated to Alexa Fluor 488	rabbit anti-jellyfish	with GFP	1:500	Clontech
D	goat anti-mouse Alexa Fluor 568	secondary pre-conjugated to Alexa Fluor 568	goat anti-mouse	with mouse antibodies	1:500	Invitrogen
E	anti-trypanosome	primary	mouse anti-trypanosome	should not be present	1:4	Dr. Pearson

with 500 µl PBS. Cells were resuspended in 200 µl PBS and subjected to 4 freeze-thaw cycles.

Protein dot blots utilizing a 1:1000 dilution of anti-GFP antibody pre-conjugated to horse radish peroxidase (Clontech, Palo Alto, CA) were performed and visualized as stated in Chapter 2.

3.3 Results

3.3.1 Plasmid Construction and Bacterial Transformation

The following plasmids were successfully used to transform *TOP10F'* cells: pEGFP-N1, pGBA1·EGFP, and pGBA2·EGFP. Figure 3.2 shows bands within a 0.7% agarose gel corresponding to true positive clones for each plasmid following amplification during the screening procedure. Expected band sizes are as follows: 0.7 kb for pEGFP-N1, 2.4 kb for pGBA1·EGFP, and 2.3 kb for pGBA2·EGFP. Sequencing results confirmed the absence of mutations in all vectors.

3.3.2 HeLa Transfection and RNA Analysis

HeLa cells were successfully transfected with each of the aforementioned plasmids.

RT-PCR results indicated the presence of the expected 0.7 kb transcript in each of the transfected HeLa cultures. Negative controls showed no amplification (Figure 3.3).

Northern blotting revealed bands of cross-reactivity of expected sizes representing pEGFP-N1 (0.7 kb), pGBA1·EGFP (2.4 kb), and pGBA2·EGFP (2.3 kb; Figure 3.4).

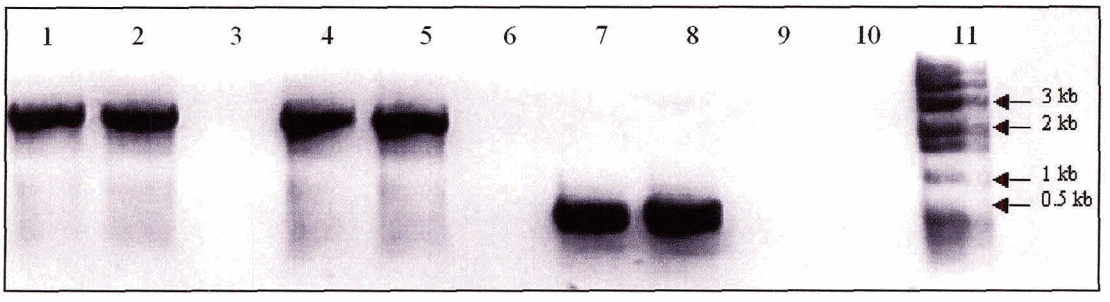


Figure 3.2. Positive clones from transformed *TOP10F'* cells screened by PCR-amplification and subsequent electrophoresis in a 0.7% agarose gel. From left to right (lanes 1-11): pGBA1•EGFP, pGBA1•EGFP, negative clone, pGBA2•EGFP, pGBA2•EGFP, negative clone, pEGFP-N1, pEGFP-N1, negative clone, negative (no DNA) control and lambda ladder. Abbreviations are as follows: EGFP (enhanced green fluorescent protein), GBA (glucocerebrosidase).

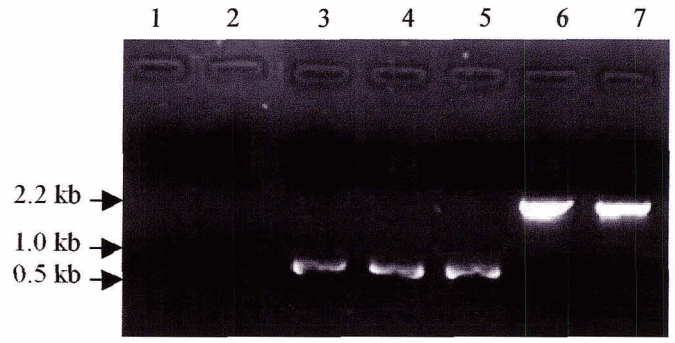


Figure 3.3. Glucocerebrosidase (GBA) and enhanced green fluorescent protein (EGFP) RNA isolated from HeLa cells at 48 hours post-transfection, subjected to RT-PCR and electrophoresed in a 0.7% agarose gel. RNA from lanes 1-5 was amplified using primers N and P, while RNA from lanes 6-7 was amplified using primers M and P. From left to right (lanes 1-7): negative control (no reverse transcriptase added), HeLa cells, pEGFP-N1, pGBA1•EGFP, pGBA2•EGFP, pGBA1•EGFP, pGBA2•EGFP.

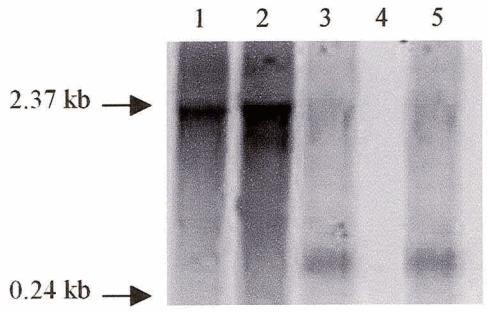


Figure 3.4. Northern blot of glucocerebrosidase (GBA) and enhanced green fluorescent protein (EGFP) RNA extracted at 48 hours post-transfection from HeLa cells. From left to right (lanes 1-5): pGBA1•EGFP, pGBA2•EGFP, pEGFP, HeLa cells, and EGFP RNA control.

Taken together, RT-PCR and Northern blot results indicate successful transcription of all constructs.

3.3.3 Microscopic Analysis and Antibody Labelling

Green fluorescence was regularly observed in pEGFP-N1-expressing cultures at each time point following 24 hours post-transfection. No fluorescence was observed at any time in control HeLa cells. Consistently fewer cells were observed to be fluorescing for each of the cultures expressing glucocerebrosidase-containing plasmids at 48 hours and 72 hours post-transfection (Figure 3.5). Treatment with trypan blue and observation of cell morphology under white light conditions indicated that these fluorescing cells were viable.

Initial examination of double-labelled cells did not indicate co-localization of either glucocerebrosidase construct (pGBA1·EGFP and pGBA2·EGFP) with the organelle-marking antibodies A and B (Figure 3.6). Fluorescence indicative of possible lysosomal positioning with exclusion of nuclear localization was present, however, low levels of fluorescence resulted in hazy images, making accurate resolution of placement difficult. Since lysosomes are distributed throughout the cytoplasm, high levels of fluorescence are required to distinguish between lysosomal localization and cytoplasmic localization. The pEGFP-N1-expressing cells demonstrated a cytoplasmically-positioned signal (Figure 3.6).

In the blockage studies, pEGFP-N1-expressing cells indicated fluorescence localized to the cytoplasm. Fluorescent patterns seemed more diffuse due to disruption of regular cell organization (Figure 3.7). For pGBA1·EGFP-transfected cells, the low levels

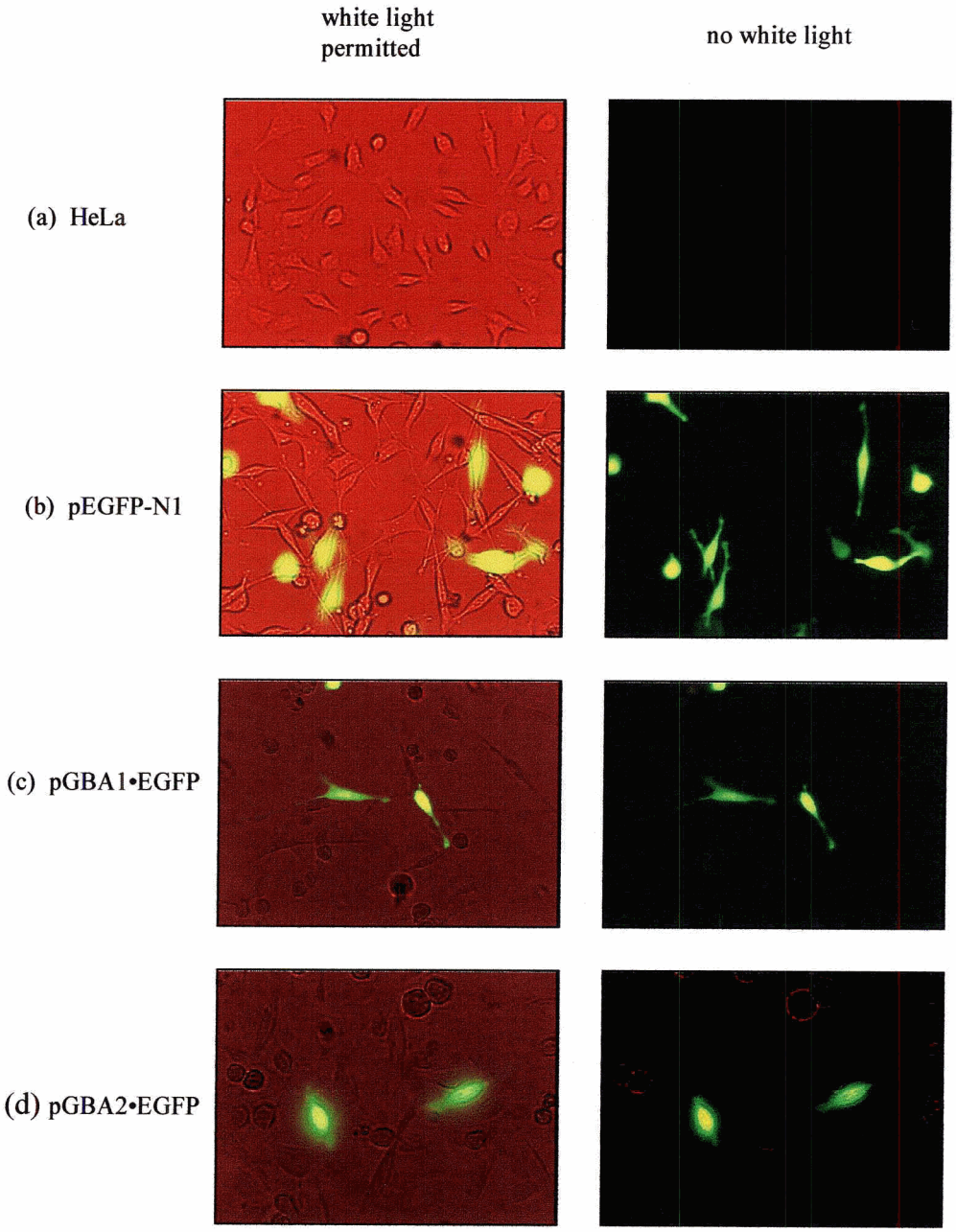


Figure 3.5. Glucocerebrosidase (GBA) and enhanced green fluorescent protein (EGFP) expression in HeLa cells 72 hours post-transfection.

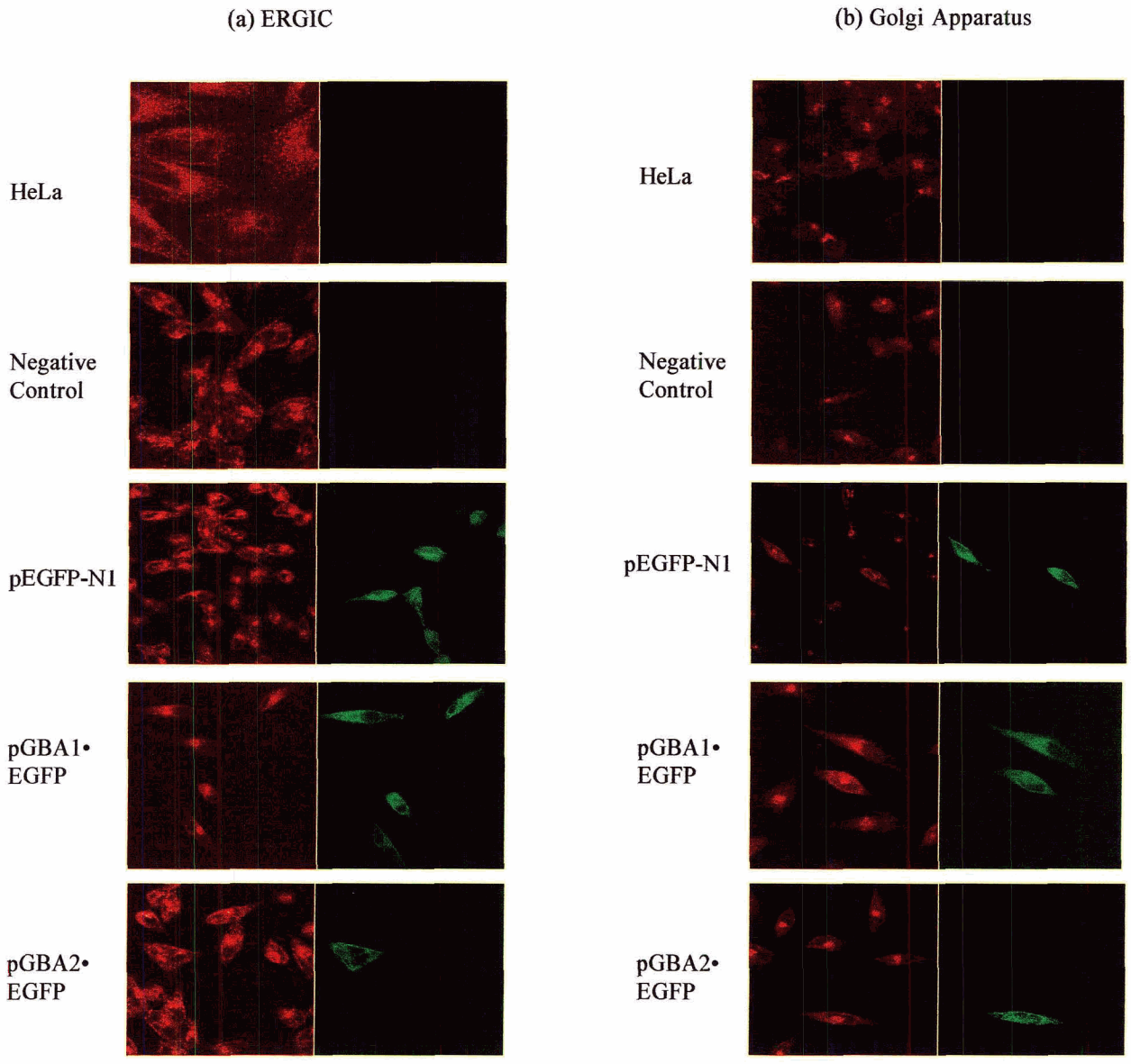


Figure 3.6. Localization of glucocerebrosidase and enhanced green fluorescent protein in HeLa cells 72 hours post-transfection. With the exception of negative controls, immunodetection was performed in (a) using primary antibodies B and C and secondary antibody D. In (b), primary antibodies A and C and secondary antibody D were used. For negative controls, primary antibody E and secondary antibody D were used. Abbreviations are as follows: ERGIC (endoplasmic reticulum-Golgi intermediate compartment), EGFP (enhanced green fluorescent protein), GBA (glucocerebrosidase). All photographs were taken at 40X magnification except for the HeLa photo in (a), which was taken at 63X magnification.

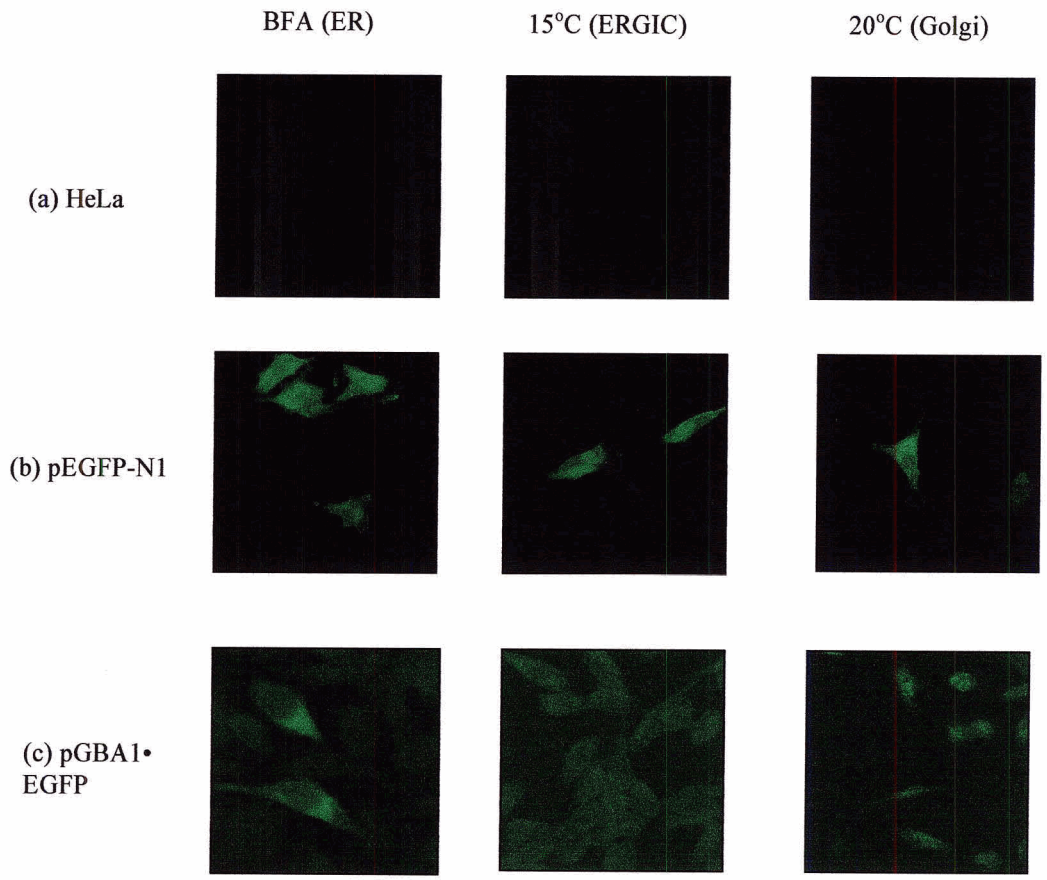


Figure 3.7. Expression of glucocerebrosidase and enhanced green fluorescent protein in HeLa cells 72 hours post-transfection after installation of transport blocks. Abbreviations are as follows: BFA (brefeldin A), ER (endoplasmic reticulum), ERGIC (endoplasmic reticulum-Golgi intermediate compartment), EGFP (enhanced green fluorescent protein), GBA (glucocerebrosidase).

of fluorescence visualized in unblocked cells (Figure 3.6) diminished significantly, indicating possible degradation of the blocked chimerae (Figure 3.7) In the case of incubation with brefeldin A, pGBA1·EGFP exhibited some localization with endoplasmic reticulum structures (Figure 3.7).

3.3.4 Protein Isolation and Analysis

Protein dot blots utilizing the GFP antibody demonstrated cross-reactivity with pEGFP-N1-expressing cells 72 hours post-transfection. No other transfected cultures showed cross-reactivity. Negative controls were also devoid of cross-reactivity (Figure 3.8). This indicates unsuccessful or inefficient translation of all glucocerebrosidase-containing constructs, or translation at levels below detection.

3.4 Discussion

Due to the search for more efficient and cost-effective methods of glucocerebrosidase production for the treatment of Gaucher disease, more attention has been focused on enzyme biosynthesis and trafficking. Though expression of glucocerebrosidase has been attempted in various cell types (Cramer et. el. 1996, Pasmanik-Chor et. al. 1996, Liu et. al. 1998), many questions still remain unanswered with respect to synthesis and trafficking. For example, resultant low levels of glucocerebrosidase production in some heterologous protein expression systems have not been explained. Another such topic in question is the mechanism of movement from the *trans* Golgi network to the lysosome. Glucocerebrosidase appears to utilize neither the mannose-6-phosphate targeting system, nor a cytoplasmic tail targeting mechanism

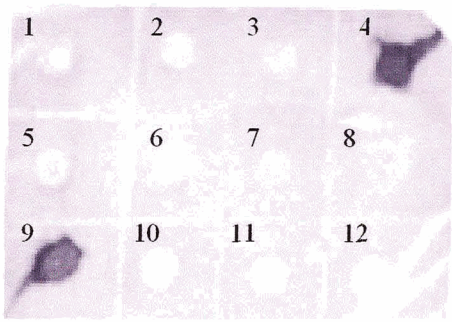


Figure 3.8. Protein dot blot of glucocerebrosidase (GBA) and enhanced green fluorescent protein (EGFP) crude lysate from transiently transfected HeLa cells. Immunodetection was performed using an anti-GFP antibody. Boxes 1-3 and 7-8 utilized crude lysate 24 hours post-transfection. Boxes 5-6 and 9-12 utilized crude lysate from 72 hours post-transfection. At each time point, lysate from two cultures expressing different clones of each glucocerebrosidase construct was examined. Boxes 1-12 are as follows: (1) pGBA1•EGFP (2) pGBA1•EGFP (3) HeLa (4) GFP positive control (5) pGBA1• EGFP (6) pGBA1•EGFP (7) pGBA2•EGFP (8) pGBA2•EGFP (9) pEGFP-N1 (10) HeLa (11) pGBA2•EGFP (12) pGBA2•EGFP.

(Beutler and Grabowski 2001). Thus, in order to examine glucocerebrosidase biosynthesis *in vivo* in a higher eukaryotic system and to search for the elusive trafficking mechanism, glucocerebrosidase/enhanced green fluorescent protein (EGFP) fusion vectors were introduced into HeLa cells. Results indicated successful transcription of both EGFP-tagged glucocerebrosidase constructs (pGBA1·EGFP and pGBA2·EGFP), as well as of EGFP alone. RT-PCR revealed the presence of transcript in each culture, with mock-transfected HeLa negative control cultures showing no amplification. Northern blotting showed bands of cross-reactivity of expected sizes for each transcript. Negative controls did not demonstrate cross-reactivity. All RNA analyses therefore confirmed the presence of transcript in all expected cases. Such successful transcription of glucocerebrosidase has been demonstrated in other studies (Pasmanik-Chor et. al. 1996, Xu and Grabowski 1998), though results from this thesis represent the first confirmation of successful glucocerebrosidase/EGFP transcription.

The results from protein analyses were less convincing. Green fluorescence was regularly observed in EGFP-expressing cultures, with no fluorescence present in control HeLa cells. This confirmed the reliability and integrity of the chosen DNA delivery system. With respect to glucocerebrosidase, however, consistently fewer cells were observed to be fluorescing, indicating low levels of translation. These results were corroborated with protein dot blots. EGFP-expressing cultures showed cross-reactivity, while cultures expressing either form of glucocerebrosidase/EGFP had expression levels below detection. One other study has reported successful expression of a glucocerebrosidase/GFP construct (Shimizu et. al. 1998). In this case, a retrovirus containing glucocerebrosidase fused to GFP was introduced into the packaging cell line

GP+envAM12. Successful expression was measured by GFP fluorescence and glucocerebrosidase activity assays. No western blots or RNA analyses were performed. Unfortunately, no further articles utilizing the created construct have been published, hindering further comparisons.

In double-labelling experiments, none of the proteins in the construct-expressing cultures co-localized with any of the organelle markers. EGFP, a cytoplasmic protein, showed localization in the cytoplasm. No vesicular-shaped pockets of concentrated fluorescence indicative of lysosomal positioning were present. With respect to glucocerebrosidase-expressing cells, a slightly speckled pattern of fluorescence hinted at some possible lysosomal localization, however, a greater intensity of fluorescence would be required to differentiate between cytoplasmic and lysosomal positioning.

In order to arrest glucocerebrosidase at different stages of movement through the biosynthetic pathway, temperature and chemical blocks were employed. Previous studies have shown that, due to interference with coatomer protein I (COPI) activity, incubation with BFA specifically and reversibly blocks translocation of proteins from the endoplasmic reticulum (ER) to the Golgi apparatus without affecting endocytosis or lysosome function (Misumi et. al. 1986, Lippincott-Schwartz et. al. 1989, Scales et. al. 1997, Ward et. al. 2001). Further along the biosynthetic pathway, incubation at 15°C has been demonstrated to block proteins in the endoplasmic reticulum-Golgi intermediate compartment (ERGIC), likely due to a shift in motor protein activity (Saraste and Kuismanen 1984, Presley et. al. 1997, Klumperman et. al. 1998, Blum et. al. 2000, Hauri et. al. 2000). Finally, incubation at 20 °C has been demonstrated to block proteins in the *trans* Golgi network (TGN; Brown et. al. 1995, Wacker et. al. 1997, Farquhar and Palade

1998, Volchuk et. al. 2000). No explanation for the mechanism of blockage at 20°C has been described (K. Howell, pers. comm.). Under each of these three conditions, EGFP appeared to localize to the cytoplasm, with a generally more diffuse pattern of fluorescence than in the unblocked situation. For pGBA1·EGFP-expressing cells, fluorescence intensity significantly diminished, indicating possible degradation of blocked chimerae. Upon incubation with BFA, pGBA1·EGFP exhibited some localization to ER structures. These results, however, underscore the need for high levels of expression in order to accurately determine localization of the desired protein.

In the case of glucocerebrosidase, successful transcription was evident from RNA analyses, but low levels of translation were apparent, impairing proper assessment of localization. Other studies in mammalian cells have also reported low levels of translation. In one example, Xu and Grabowski (1998) estimated the efficiency of glucocerebrosidase mRNA translation by the ratio of mRNA molecules per cell to the protein mass in picograms. This ratio in untransduced human fibroblasts was 200, while that in transduced human fibroblasts was 26,991, indicating a 135-fold less efficient usage of glucocerebrosidase mRNA in transduced cells. In transfected CHO cells, transduced Gaucher disease fibroblasts, and transduced mouse C2C12 cells, there was between 55- and 98-fold less efficient usage of glucocerebrosidase mRNA than in untransduced counterparts. In mammalian cells, this translational inefficiency has been proposed to result from the interaction of the glucocerebrosidase transcript with an 80,000 MW mRNA-binding translational control protein (TCP80; Xu and Grabowski 1998, Xu et. al. 2000). TCP80 has been shown to specifically bind a 184 nucleotide region (261 bp - 444 bp) of glucocerebrosidase mRNA, preventing interaction of the

mRNA with polysomes and subsequent translation initiation (Xu and Grabowski 1999). This interaction is further examined in Chapter 4.

3.5 References

Aerts, J.M.F.G., Brul, S., Donker-Koopman, W.E., van Weely, S., Murray, G.J., Barranger, J.A., Tager, J.M., and Schram, A.W. (1986). Efficient routing of glucocerebrosidase to lysosomes requires complex oligosaccharide chain formation. *Biochem. Biophys. Res. Comm.* 141: 452-458.

Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K., and Watson, J.D. (1994). *Molecular Biology of the Cell* (3rd ed.). New York: Garland.

Beutler, E. and Grabowski, GA. (2001). Gaucher disease. In Scriver, C.R., Beaudet, A.L, Sly, W.S., and Valle, D (eds.), *The Metabolic and Molecular Bases of Inherited Disease* (vol. 3, 8th ed., pp. 3635-3668). New York: McGraw-Hill.

Blum, R., Stephens, D.J., and Schulz, I. (2000). Lumenal targeted GFP, used as a marker of soluble cargo, visualizes rapid ERGIC to Golgi traffic by a tubulo-vesicular network. *J. Cell Sci.* 113: 3151-3159.

Brady, R.O. (1966). The sphingolipidoses. *N. Engl. J. Med.* 275: 312-318.

Brown, W.J., DeWald, D.B., Emr, S.D., Plutner, H., and Balch, W.E. (1995). Role for phophatidylinositol 3-kinase in the sorting and transport of newly synthesized lysosomal enzymes in mammalian cells. *J. Cell Biol.* 130: 781-796.

Campbell, T.N. and Choy, F.Y.M. (2001). Protein trafficking in the biosynthetic pathway. *Molecular Biology Today* 2: 67-76.

Cramer, C.L., Weissenborn, D.L., Oishi, K.K., Grabau, E.A., Bennett, S., Ponce, E., Grabowski, G.A., and Randin, D.N. (1996). Bioproduction of human enzymes in transgenic tobacco. *Ann. New York Acad. Sci.* 792:62-71.

de Fost, M., Aerts, J.M.F.G., and Hollak, C.E.M. (2003). Gaucher disease: from fundamental research to effective therapeutic interventions. *Neth. J. Med.* 61: 3-8.

Farquhar, M.G. and Palade, G.E. (1998). The Golgi apparatus: 100 years of progress and controversy. *Trends Cell Biol.* 8: 2-10.

Glickman, J.N. and Kornfield, S. (1993). Mannose 6-phosphate-independent targeting of lysosomal enzymes in I-cell disease B lymphoblasts. *J. Cell. Biol.* 123: 99-108.

Gussow, D. and Clackson, T. (1989). Direct clone characterization from plaques and colonies by the polymerase chain reaction. *Nucleic Acids Res.* 17: 4000.

Hauri, H-P., Kappeler, F., Andersson, H., and Appenzeller, C. (2000). ERGIC-53 and traffic in the secretory pathway. *J. Cell Sci.* 113: 587-596.

Klumperman, J., Schweizer, A., Clausen, H., Tang, B.L., Hong, W., Oorschot, V., and Hauri, H-P. (1998). The recycling pathway of protein ERGIC-53 and dynamics of the ER-Golgi intermediate compartment. *J. Cell Sci.* 111: 3411-3425.

Lehmann, L.E., Eberle, W., Krull, S., Prill, V., Schmidt, B., Sander, C., von Figura, K., and Peters, C. (1992). The internalization signal in the cytoplasmic tail of lysosomal acid phosphatase consists of the hexapeptide PGYRHV. *EMBO J.* 11: 4391-4399.

Lemansky, P., Gieselmann, V., Hasilik, A., and von Figura, K. (1985). Synthesis and transport of lysosomal acid phosphatase in normal and I-cell fibroblasts. *J. Biol. Chem.* 260: 9023-9030.

Leonova, T., and Grabowski, G.A. (2000). Fate and sorting of acid ß-glucosidase in transgenic mammalian cells. *Mol. Genet. Metabol.* 70: 281-294.

Lippincott-Schwartz, J., Yuan, L.C., Bonifacino, J.S., and Klausner, R.D. (1989). Rapid redistribution of Golgi proteins into the ER in cells treated with brefeldin A: evidence for membrane cycling from the Golgi to ER. *Cell* 56: 801-813.

Liu, C., Bahnson, A.B., Dunigan, J.T., Watkins, S.C., and Barranger, J.A. (1998). Long-term expression and secretion of human glucocerebrosidase by primary murine and human myoblasts and differentiated myotubes. *J. Mol. Med.* 76: 773-781.

Misumi, Y., Misumi, Y., Miki, K., Takatsuki, A., Tamura, G., and Ikehara, T. (1986). Novel blockade by brefeldin A of intracellular transport of secretory proteins in cultured rat hepatocytes. *J. Cell Biochem.* 261: 11398-11403.

Pasmanik-Chor, M., Elroy-Stein, O., Aerts, H., Agmon, V., Gatt, S., and Horowitz, M. (1996). Overexpression of human glucocerebrosidase containing different-sized leaders. *Biochem. J.* 317: 81-88.

Presley, J.F., Cole, N.B., Schroer, T.A, Hirschberg, K., Zaal, K.J.M., and Lippincott-Schwartz, J. (1997). ER-to-Golgi transport visualized in living cells. *Nature* 389: 81-85.

Saraste, J. and Kuismanen, E. (1984). Pre- and post-Golgi vacuoles operate in the transport of Semliki Forest virus membrane glycoproteins to the cell surface. *Cell* 38: 535-549.

Scales, S.J., Pepperkok, R., and Kreis, T.E. (1997). Visualization of ER-to-Golgi transport in living cells reveals a sequential mode of action for COPII and COPI. *Cell* 90:1137-1148.

Shimizu, T., Ando, K., Kimura, M., Miyatake, H., Inokuchi, S., Takakura, I., Migita, M., Shimada, T. and Kato, S. (1998). A simple and efficient purification of transduced cells by using green fluorescent protein gene as a selection marker. *Acta Paediatr. Jpn.* 40: 586-592.

Sorge, J.A., West, C., Kuhl, W., Treger, L., and Beutler, E. (1987). The human glucocerebrosidase gene has two functional ATG initiator codons. *Am. J. Hum. Genet.* 41: 1016-1024.

Volchuk, A., Amherdt, M., Ravazzola, M., Brugger, B., Rivera, V.M., Clackson, T., Perrelet, A., Sollner, T.H., Rothman, J.E., and Orci, L. (2000). Megavesicles implicated in the rapid transport of intracisternal aggregates across the Golgi stack. *Cell* 102: 335-348.

Wacker, I., Kaether, C., Kromer, A., Migala, A., Almers, W., and Gerdes, H-H. (1997). Microtubule-dependent transport of secretory vesicles visualized in real time with a GFP-tagged secretory protein. *J. Cell Sci.* 110: 1453-1463.

Ward, T.H., Polishchuk, R.S., Caplan, S., Hirschberg, K., and Lippincott-Schwartz, J. (2001). Maintenance of Golgi structure and function depends on the integrity of ER export. *J. Cell. Biol.* 155: 557-570.

Xu, Y-H., Busald, C., and Grabowski, G.A. (2000). Reconstitution of TCP80/NF90 translation inhibition activity in insect cells. *Mol. Genet. Metab.* 70: 106-115.

Xu, Y-H. and Grabowski, G.A. (1998). Translational inefficiency of acid β-glucosidase mRNA in transgenic mammalian cells. *Mol. Genet. Metab.* 64: 87-98.

Xu, Y-H. and Grabowski, G.A. (1999). Molecular cloning and characterization of a translational inhibitory protein that binds to coding sequences of human acid β-glucosidase and other mRNAs. *Mol. Genet. Metab.* 68: 441-454.

Chapter 4 – Examination of Glucocerebrosidase and Enhanced Green Fluorescent Protein Biosynthesis in HeLa and COS-1 Cells Through Truncated Constructs and siRNA Interference

4.1 Introduction

One issue of particular interest in the search for improved Gaucher disease treatment is why an apparent translational inefficiency occurs when glucocerebrosidase is overexpressed in some heterologous protein production systems. While all such systems utilize highly efficient transcriptional promoters and/or high gene copy strategies to maximize transcription, this has not resulted in high protein expression levels. This phenomenon was evident in Chapter 2 of this dissertation. Similarly, the commercial product imiglucerase, which is produced in a modified Chinese Hamster Ovary (CHO) system, is faced with higher than expected production costs due to current limitations in enzyme yield (Mistry et. al. 1996). In another example, Xu and Grabowski (1998) noted that there was between 55- and 98-fold less efficient usage of mRNA in transfected CHO cells, transduced Gaucher disease fibroblasts, and transduced mouse C2C12 cells, than in untransduced/untransfected counterparts. In mammalian cells, this translational inefficiency has been proposed to result from the interaction of the glucocerebrosidase transcript with an mRNA-binding protein labelled TCP80 (Xu and Grabowski 1998, Xu et. al. 2000). The cDNA sequence of TCP80 is 96% similar to cyclosporin FK506-sensitive nuclear factor (NF90) and 99% similar to M phase phosphoprotein (MPP4). In some instances, these proteins have been referred to as the same entity (Duchange et. al. 2000, Xu et. al. 2000). TCP80 has been shown to specifically bind a 184 nucleotide

region (261 bp - 444 bp) of glucocerebrosidase mRNA, thus preventing interaction of the mRNA with polysomes and subsequent translation initiation (Xu and Grabowski 1999). TCP80 has also been demonstrated to be phosphorylated by the classical protein kinase C pathway, and deficient phosphorylation of TCP80 facilitates glucocerebrosidase translation inhibition (Zhao and Grabowski 2002).

As noted in Chapter 3 of this thesis, such translational inefficiency of glucocerebrosidase appeared to be present in HeLa cells. Since TCP80 has been proposed to contribute to this inefficiency, it follows that removal or blockage of TCP80 could lead to increased translational levels of glucocerebrosidase. In order to test this hypothesis, two approaches were taken. In the first approach, two truncated constructs were created. One was an EGFP-tagged 500 bp segment of glucocerebrosidase containing the proposed 184 nucleotide TCP80-binding site. The second was an EGFP-tagged 500 bp segment of glucocerebrosidase downstream of the 184 nucleotide binding site. Expression of these constructs would therefore aid in elucidating whether the proposed binding site contributes to glucocerebrosidase translational inhibition.

In the second approach, small interfering RNAs (siRNAs) were employed. The discovery and manipulation of RNA-mediated interference (RNAi), a term coined by Fire et. al. (1998), has presented the scientific community with new ways to examine old mechanisms. RNAi is a multi-step process involving the generation of siRNAs *in vivo* through the action of the RNase III endonuclease Dicer. The resulting 21-23 nucleotide siRNAs mediate degradation of their complementary RNA (Shi 2003; Figure 4.1). Since the initial discovery, RNAi has been used to selectively knock down specific genes in organisms ranging from plants (Chuang and Megerowitz 2000) to worms (Fire et. al.

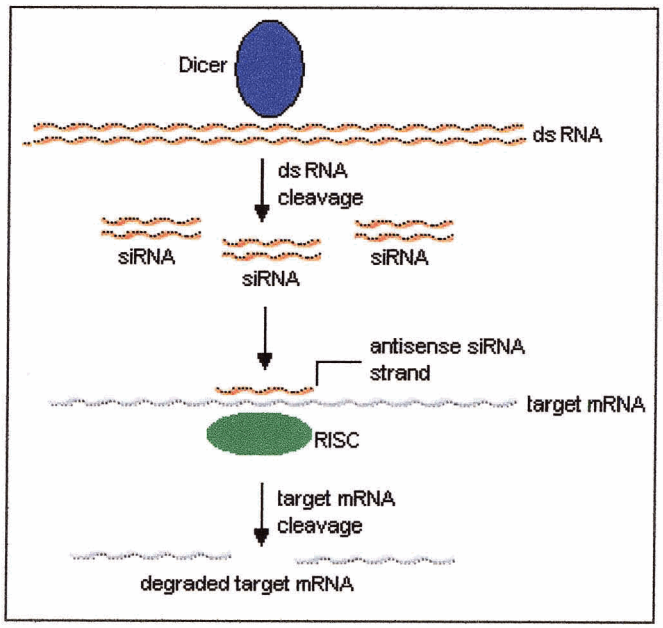


Figure 4.1. Schematic diagram of small interfering RNA (siRNA) synthesis and action. Long double-stranded RNA (dsRNA) is processed by Dicer to produce siRNAs. The antisense strand of the siRNA is used by an RNA-induced silencing complex (RISC) to guide mRNA cleavage, promoting mRNA degradation. This figure has been modified from McManus and Sharp (2002).

1998) to mammals (Elbashir et. al. 2002, Scherr et. al. 2003). Therefore, in order to test the theory of TCP80 inhibition contributing to the incongruence of glucocerebrosidase transcriptional and translational levels, an siRNA to knock down TCP80 was designed. A control siRNA and a GFP-directed siRNA were also utilized to test system integrity. Expression of EGFP-tagged glucocerebrosidase constructs beginning at either the first or the second ATG initiation site was examined in co-transfection experiments with each type of siRNA.

4.2 Methods and Materials

4.2.1 HeLa and COS-1 Culture

HeLa cells (donation from Dr. T. Pfeiffer) were cultured at 37°C with 5% CO_2 in RPMI 1640 with 10% fetal bovine serum and 5% penicillin/streptomycin (Invitrogen, Carlsbad, CA). COS-1 cells (donation from Dr. K. McNagny) were cultured at 37 °C with 5% CO_2 in DMEM with 10 % fetal bovine serum, 5% penicillin/streptomycin, and 4 *mM* L-glutamine (Invitrogen, Carlsbad, CA).

4.2.2 Plasmid Construction and Bacterial Transformation

Construction of plasmids pGBA1·EGFP (full-length glucocerebrosidase) and pGBA2·EGFP (glucocerebrosidase starting from the second ATG initiation site) is described in Chapter 3.

To create the truncated 500 bp segment of glucocerebrosidase containing the proposed 184 bp inhibitor-binding site (pGBA3·EGFP), glucocerebrosidase template was amplified with *Eco* RI restriction site-containing primers Q and R, digested, purified and

cloned into pEGFP-N1. Primer Q also included an in-frame initiator ATG codon. The truncated 500 bp segment at the 3' end of glucocerebrosidase devoid of the 184 bp binding site was amplified with *Eco* RI restriction site-containing primers S and L prior to creation of pGBA4·EGFP in the same manner as pGBA3·EGFP. Primer S also included an in-frame initiator ATG codon. All primer information is contained in Table 4.1

Bacterial transformation, screening, and large-scale isolation were performed as described in Chapters 2-3 using primers Q and R for pGBA3·EGFP and primers S and L for pGBA4·EGFP.

4.2.3 HeLa Transfection with Truncated Plasmids

HeLa cells in 6-well plates were transfected at 80% confluency with FuGENE6 in a 3:1 ratio (μl FuGENE6: μg DNA) according to the manufacturer's instructions (Roche, Indianapolis, IN). For mock-transfected HeLa culture controls, FuGENE6 was added alone. Fetal bovine serum (10%) was added to the wells 3-5 hours post-transfection. In cultures set aside for microscopic visualization, four sterile 12mm round coverslips were coated with poly-lysine, dried overnight, and placed in the bottom of each well prior to seeding with trypsinized HeLa cells. Cells were grown to 80% confluency and similarly transfected with FuGENE6.

4.2.4 siRNA Design

All target sequences and siRNA sequences are shown in Table 4.2. The sequence for TCP80 was obtained from GenBank (accession number AF141870). A BLAST search

Table 4.1. Primers used for glucocerebrosidase (GBA) and enhanced green fluorescent protein (EGFP) plasmid construction and RNA analysis.

Primer	**DNA**	**Sequence (5' to 3')**	**Location in cDNA[a]**	**Orientation**
L	GBA	ATCGAATTCGTCCCTGGCGATGCC AC	1621-1596	antisense
M	GBA	CCTGTGAGGCTGCCAGCCATGAA TTCGGTA	51-80	sense
N	EGFP	ATGGTGAGCAAGGGCGAGGAGCT GT	679-703	sense
P	EGFP	TACAGCTCGTCCATGCCGAGAGT GATCC	1391-1364	antisense
Q	GBA	CGAATTCGGATGCGGATGGAGCT GAGTATGGG	247-278	sense
R	GBA	CCAGAATTCTCACAAAGTATCTGG CCCAGGTCTG	769-736	antisense
S	GBA	AACGAATTCTGATGCCCAACACC ATGCTCTTT	1097-1128	sense
T	TCP80	TGCGTCCAATGCGAATTTTTGTGA ATG	113-139	sense
U	TCP80	TGGATGGAACGGTTGTCGGCC	514-494	antisense

[a] The first base of the upstream initiator codon of glucocerebrosidase is designated #1 (Sorge et. al. 1987). Numbers for EGFP correspond to those in Clontech manuals (Palo Alto, CA). Numbers for TCP80 correspond to those in GenBank.

Note. No primer was designated "O" since this may be confused with zero in the text of this thesis.

Table 4.2. Target and small interfering RNA (siRNA) sequences used in co-transfection experiments with glucocerebrosidase and green fluorescent protein (GFP).

Gene	**Organism**	**Target Sequence**	**Location in Target Sequence**[a]	**siRNA Sequence**[b]
TCP80	Human	CACGTGGCCG TTAAGGTGTT A	1477-1497	CGUGGCCGUUAA GGUGUUAUU (s) UAACACCUUAAC GGCCACGUG (a)
GFP	*Aequorea victoria*	CGGCAAGCTG ACCCTGCCGTT CAT	122-143	GCAAGCUGACCC UGAAGUUCAU (s) GAACUUCAGGGU CAGCUUGCCG (a)
None (section 21 of 136 of the complete genome)	*Thermotoga meritimia*	AATTCTCCGA ACGTGTCACG T	section 21 of 136	UUCUCCGAACGU GUCACGUTT (s) ACGUGACACGUU CGGAGAATT (a)

[a]In the TCP80 sequence, 1 represents the first base in the GenBank sequence. For GFP, 1 represents the first nucleotide of the start codon.

[b]For siRNA sequences, (s) = sense and (a) = antisense.

was performed to test for homologous sequences. Near-identical matches included NF90, MPP4, DRBP76 (double-stranded RNA-binding nuclear protein) and ILF3 (interleukin enhancer binding factor 3). It has been proposed that some or all of these may actually be the same gene/protein (Duchange et. al. 2000, Xu et. al. 2000).

Two online siRNA design programs were utilized: one from the Dharmacon siDESIGN Center (Dharmacon, Lafayette, CO) and one from the Xeragon/QIAGEN siRNA Design Site (QIAGEN, Valencia, CA). The TCP80 sequence was analyzed by both programs and numerous possible target sequences were presented. The field was narrowed down to only those target sequences within the proposed RNA binding sites of TCP80. Since none of the proposed targets within these sites followed the current rules of siRNA design, modifications were made. Current considerations for selecting target sites for siRNA design include: choosing a 21-23 nucleotide sequence in the coding mRNA with a GC ratio close to 50%, avoiding more than 3 G's in a row, choosing sequences starting with AA or CA, and ensuring that the target sequence is not homologous to other genes (QIAGEN, Valencia, CA). The final target site (and corresponding siRNA sequence) for TCP80 is in Table 4.2. A BLAST search of this target sequence was performed. No cases of homology, other than the previously mentioned NF90/MPP4/DRBP76/ILF3, were encountered.

A 21 bp siRNA from *Thermotoga maritimia*, purchased from QIAGEN (Valencia, CA), was utilized as a control. The target sequence, based on a BLAST search, did not correspond to any gene, but instead represented section 21 of 136 of the complete *T. maritimia* genome (Table 4.2).

A 22 bp siRNA, purchased from QIAGEN (Valencia, CA), was used as a

positive control (Table 4.2). This siRNA was previously determined to knock down GFP expression in mammalian systems without affecting other genes (Caplen et. al. 2001).

4.2.5 Co-transfection of HeLa and COS-1 Cells

Initial attempts at co-transfecting with FuGENE6 resulted in successful transcription of introduced DNA, as noted in Northern blots (results not shown). Fluorescence of pEGFP-N1 was also present 24-48 hours post-transfection, indicating successful expression. However, no successful transfer of siRNA into cells was apparent, as noted by the presence of similar levels of RNA and fluorescence in the EGFP/control-siRNA and EGFP/EGFP-siRNA transfected cells. Therefore, a new transfection reagent, TransMessenger™ (QIAGEN, Valencia, CA), was employed.

Transfection with TransMessenger™ was modified from the manufacturer's protocol. Separate 12-well plates were seeded with HeLa and COS-1 cells and grown to 80% confluency over 24-30 hours. On the day of transfection, 4.8 µl of Enhancer R was diluted with 94 µl of Buffer EC-R. Next, 0.4 µg of plasmid DNA (pEGFP-N1, pGBA1·EGFP, or pGBA2·EGFP) and 0.2 µg of siRNA (control, EGFP, or TCP80) were added. For control HeLa and COS-1 cultures, no plasmid DNA was added. This mixture was vortex-mixed, incubated for 5 minutes at room temperature, and pulsed. TransMessenger™ (2.5 µl) was then added to the siRNA/DNA/Enhancer R mixture, which was then vortex-mixed and incubated at room temperature for 10 minutes. During complex formation, medium was aspirated from the cells in the 12-well plates which were then washed with 2 ml sterile PBS. At the end of the 10 minute incubation, 200 µl serum-free and antibiotic-free medium (RPMI 1640 for HeLa cells; DMEM for COS-1

cells) was added to the nucleic acids/TransMessenger mixture which was then pipetted evenly over the cells. The cells were incubated for 3 hours under their normal growth conditions, followed by removal of the complexes, one wash with sterile PBS, and addition of serum-containing medium.

In cultures set aside for microscopic visualization, one sterile 12 mm round coverslip was placed in the bottom of each well prior to seeding with either HeLa or COS-1 cells. Cells were grown to 80% confluency and similarly transfected with TransMessenger™.

4.2.6 RNA Analysis of Truncated Plasmid Transfections

Total RNA was extracted 48 hours post-transfection with the RNeasy® Mini Kit according to the manufacturer's instructions for 6-well plates (QIAGEN, Valencia, CA).

For RT-PCR analysis, 1 μl of RNA was used as template for 36 cycles of amplification according to the recommended protocol for Superscript™ One-Step RT-PCR with Platinum®Taq (Invitrogen, Carlsbad, CA). Primers N and P (58°C anneal) were utilized for all reactions. A control with Platinum® Taq but without reverse transcriptase was included to test whether residual contaminating DNA was present.

Northern blotting was performed as reported in Chapter 3, except that 6-8 μg (instead of 6 μg) were added.

4.2.7 RNA Analysis of Co-transfected HeLa and COS-1 Cells

Total RNA was extracted 48 hours post-transfection with the RNeasy® Mini Kit (QIAGEN, Valencia, CA) according to the manufacturer's instructions for 12-well plates

with one exception: 175 µl (instead of 350 µl) of the lysis buffer were added to each well prior to detachment of cells with a rubber policeman. This crude lysate was then pooled from two replicate wells per co-transfected plasmid/siRNA or cell/siRNA into one tube prior to homogenization and subsequent extraction steps.

For RT-PCR analysis, 1-2 µl of RNA were used as template for 36 cycles of amplification according to the recommended protocol for Superscript™ One-Step RT-PCR with Platinum®Taq (Invitrogen, Carlsbad, CA). Primers T and U (58°C anneal) were used when testing for TCP80 presence. Primers N and P (58°C anneal) were utilized for all other reactions (Table 4.1). A control with Platinum® Taq but without reverse transcriptase was included to test whether residual contaminating DNA was present in the isolated RNA samples.

For Northern blots, a protocol slightly modified from that reported in Chapter 3 was used. In brief, 6 µg of each sample were added to 5X RNA loading dye (0.16% bromophenol blue, 4 *mM* EDTA [pH 8.0], 0.9 *M* formaldehyde, 3.2% formamide, 20% glycerol, 0.4X MOPS buffer [0.2 *M* 3-{N-morpholino}propanesulfonic acid, 0.08 *M* sodium acetate, 0.01 *M* EDTA; pH 7.0]). This was incubated at 65°C for 5 minutes, cooled on ice for 5 minutes, and loaded in a modified 1% agarose gel (1.0 g agarose, 1X MOPS buffer, 0.2 *M* formaldehyde, 0.1 µg/ml ethidium bromide). Samples were electrophoresed at 70V for 2 hours in RNA running buffer (1X MOPS buffer, 0.2 *M* formaldehyde). The RNA was transferred to a Hybond-N^+ nylon membrane (Amersham, Piscataway, NJ) by capillary transfer overnight at 4°C in 10X SSC (1.5 *M* sodium chloride, 0.15 *M* sodium citrate). Subsequently, the transferred RNA was cross-linked to the membrane by incubation at 80°C for 2 hours. When testing for TCP80 presence, a 400

bp cDNA probe (representing bp 113-514) was PCR-amplified using primers T and U from total RNA isolated from HeLa cells. For all other blots, a DNA probe representing the full-length EGFP cDNA was PCR-amplified using primers N and P. These probes were conjugated with alkaline phosphatase according to the AlkPhos Direct Manual (Amersham, Piscataway, NJ). Membranes were hybridized with 100-300 ng of labelled probe at 55°C overnight and washed as directed. Bands were developed by incubation with ECF chemifluorescent reagent (Amersham, Piscataway, NJ) as directed for 24-72 hours before visualizing with a Molecular Dynamics Storm 860 phosphorimager (Molecular Dynamics, Sunnyvale, CA).

4.2.8 Protein Analysis for Truncated and Co-transfected Plasmids

Every 24 hours, one coverslip per expression culture was removed, mounted on a clean slide with ProLong Antifade (Invitrogen, Carlsbad, CA) and analyzed by fluorescence microscopy. Addition of diluted trypan blue (donation from Dr. R. Burke) in a 1:1 ratio (volume trypan blue: volume cells) was included every second replicate to examine cell viability. Presence or absence of green fluorescence was recorded and photographed with both a Zeiss Epifluorescence microscope and a Zeiss LSM410 confocal microscope.

4.3 Results

4.3.1 Bacterial Transformation with Truncated Plasmids

The following plasmids were successfully used to transform *TOP10F'* cells: pEGFP-N1, pGBA3·EGFP, and pGBA4·EGFP. Figure 4.2 shows bands within a 0.7%

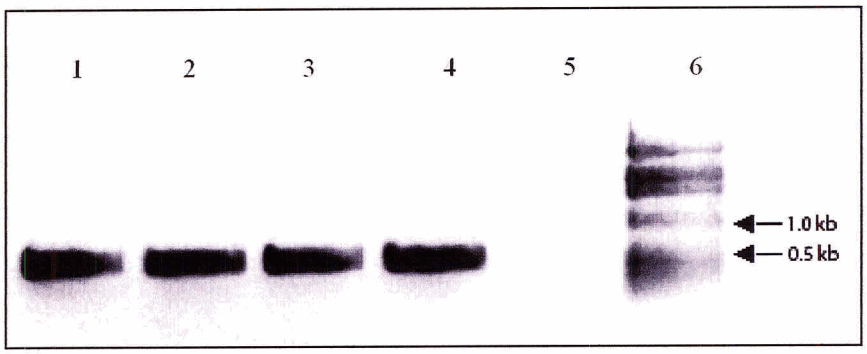


Figure 4.2. Positive clones from transformed *TOP10F'* cells screened by PCR-amplification and subsequent electrophoresis in a 0.7% agarose gel. From left to right (lanes 1-6): pGBA3•EGFP, pGBA3•EGFP, pGBA4•EGFP, pGBA4•EGFP, negative (no DNA) control, lambda ladder. Abbreviations are as follows: EGFP (enhanced green fluorescent protein), GBA (glucocerebrosidase).

agarose gel corresponding to true positive clones for each plasmid following amplification during the screening procedure. Expected band sizes are as follows: 0.5 kb for pGBA3·EGFP and 0.5 kb for pGBA4·EGFP. Sequencing results confirmed the absence of mutations in all vectors.

4.3.2 HeLa Transfection with Truncated Plasmids and RNA Analysis

HeLa cells were successfully transfected with pEGFP-N1, pGBA3·EGFP and pGBA4·EGFP.

RT-PCR results indicated the presence of the expected 0.7 kb transcript in each of the transfected HeLa cultures. Negative controls showed no amplification (Figure 4.3).

Northern blotting revealed strong bands of cross-reactivity of expected sizes representing pEGFP-N1 (0.7 kb), pGBA3·EGFP (1.2 kb), and pGBA4·EGFP (1.2 kb), with no bands present in negative controls (Figure 4.4). Taken together, RT-PCR and Northern blot results indicate successful transcription of all constructs.

4.3.3 Co-transfection of HeLa and COS-1 Cells and RNA Analysis

HeLa and COS-1 cells were successfully co-transfected in combinations of plasmid/siRNA with the following: pEGFP-N1, pGBA1·EGFP, pGBA2·EGFP, control siRNA, EGFP siRNA, and TCP80 siRNA.

In both HeLa and COS-1 cells, RT-PCR results using the TCP80 primers T and U revealed the expected 0.4 kb transcripts present for each co-transfection, with nothing visible in the negative control (Figure 4.5). Northern blotting, however, did not show any cross-reactivity, indicating RNA levels below detection (Figure 4.6). No positive control

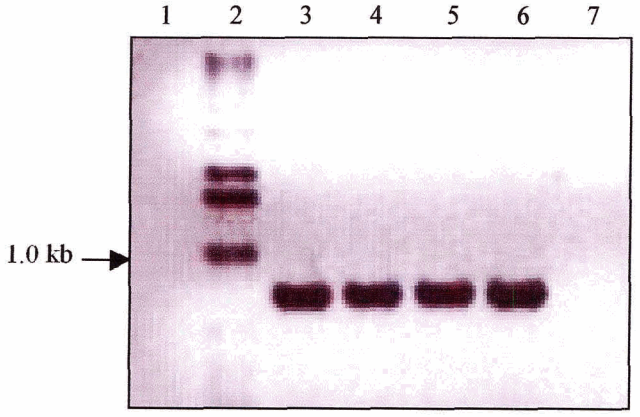


Figure 4.3. Glucocerebrosidase (GBA) and enhanced green fluorescent protein (EGFP) RNA isolated from HeLa cells 48 hours post-transfection, subjected to RT-PCR, and electrophoresed in a 1.0% agarose gel. Left to right (lanes 1-7): negative control (no reverse transcriptase), lambda ladder, pEGFP-N1 positive control, pEGFP-N1, pGBA3•EGFP, pGBA4•EGFP, HeLa cells.

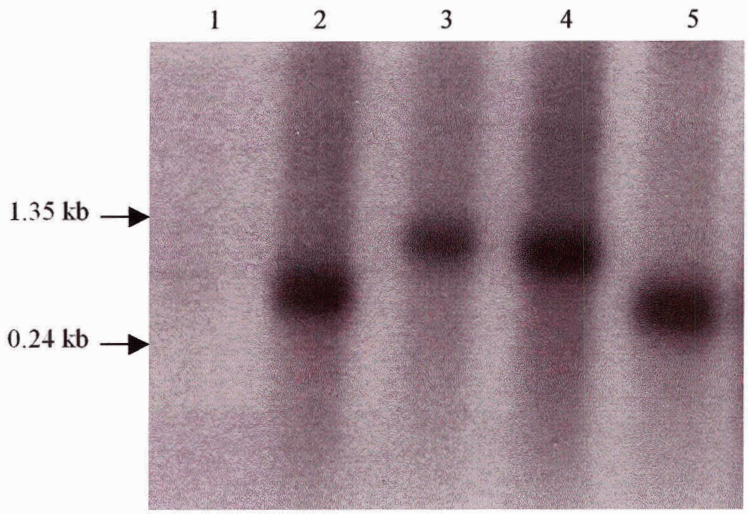


Figure 4.4. Northern blot of glucocerebrosidase (GBA) and enhanced green fluorescent protein (EGFP) RNA extracted at 48 hours post-transfection from transfected HeLa cells. From left to right (lanes 1-5): HeLa cells, EGFP RNA control, pGBA4•EGFP, pGBA3•EGFP, and pEGFP-N1.

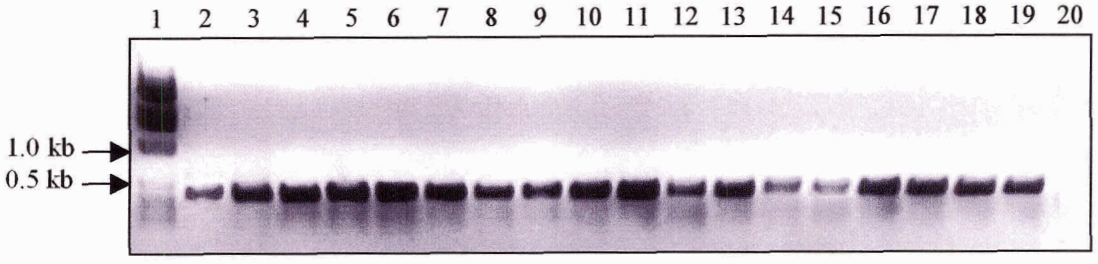


Figure 4.5. RNA isolated from HeLa (lanes 2-10) and COS-1 (lanes 11-20) cells at 48 hours after co-transfection, subjected to RT-PCR using TCP80 primers, and electrophoresed in a 0.7% agarose gel. Lanes 2-4 and 11-13 included control small interfering RNA (siRNA). Lanes 5-7 and 14-16 included EGFP siRNA. Lanes 8-10 and 17-20 included TCP80 siRNA. From left to right (lanes 1-20): lambda ladder, pEGFP-N1, pGBA1•EGFP, pGBA2•EGFP, pEGFP-N1, pGBA1•EGFP, pGBA2•EGFP, pEGFP-N1, pGBA1•EGFP, pGBA2•EGFP, pEGFP-N1, pGBA1•EGFP, pGBA2•EGFP, pEGFP-N1, pGBA1•EGFP, pGBA2•EGFP, pEGFP-N1, pGBA1•EGFP, pGBA2•EGFP, negative control (no reverse transcriptase).

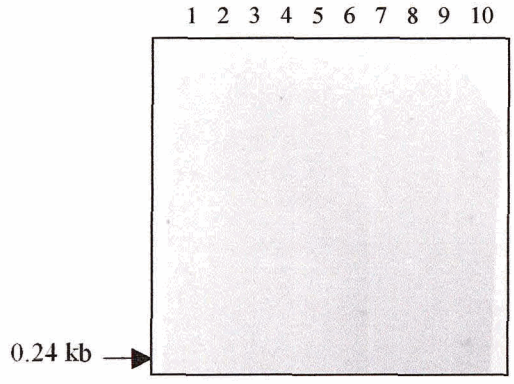


Figure 4.6. Northern blot of TCP80 RNA isolated from HeLa cells at 48 hours after co-transfection of glucocerebrosidase (GBA) and enhanced green fluorescent protein (EGFP) plasmids with small interfering RNAs (siRNAs). Lanes 1-3 included TCP80 siRNA. Lanes 4-6 included EGFP siRNA. Lanes 7-10 included control siRNA. From left to right (lanes 1-10): pGBA2•EGFP, pGBA1•EGFP, pEGFP-N1, pGBA2•EGFP, pGBA1•EGFP, pEGFP-N1, pGBA2•EGFP, pGBA1•EGFP, pEGFP-N1, negative control.

was available to test probe hybridization, however, a similar probe was used successfully by Duchange et. al. (2000).

In HeLa cells, RT-PCR results using the EGFP primers N and P indicated the presence of the expected 0.7 kb transcript in each of the transfected cultures. Negative controls showed no amplification (Figure 4.7). Northern blotting revealed bands of cross-reactivity of expected sizes representing pEGFP-N1 (0.7 kb), pGBA1·EGFP (2.4 kb) and pGBA2·EGFP (2.3 kb) in cultures co-transfected with control and TCP80 siRNA, but not in cultures co-transfected with EGFP siRNA (Figure 4.8). Thus, though RT-PCR indicated transcript presence in all cultures, RNA levels were below detection in the Northern blot of pEGFP-N1, pGBA1·EGFP and pGBA2·EGFP when co-transfected with EGFP siRNA.

In COS-1 cells, RT-PCR using EGFP primers N and P indicated the presence of the expected transcript in each of the transfected cultures except that representing pGBA2·EGFP when co-transfected with TCP80 siRNA. Negative controls showed no amplification (Figure 4.9). Northern blotting revealed bands of cross-reactivity for pEGFP-N1 when co-transfected with each form of siRNA with less intensity apparent in the EGFP siRNA co-transfected culture (Figure 4.10). Cross-reactivity was observed for both pGBA1·EGFP and pGBA2·EGFP when co-transfected with control siRNA, but only for pGBA1·EGFP when co-transfected with TCP80 siRNA. No bands were evident for pGBA1·EGFP or pGBA2·EGFP when co-transfected with EGFP siRNA (Figure 4.10).

4.3.4 Protein Analysis of Truncated Plasmids

Green fluorescence was regularly observed in pEGFP-N1-expressing cultures at

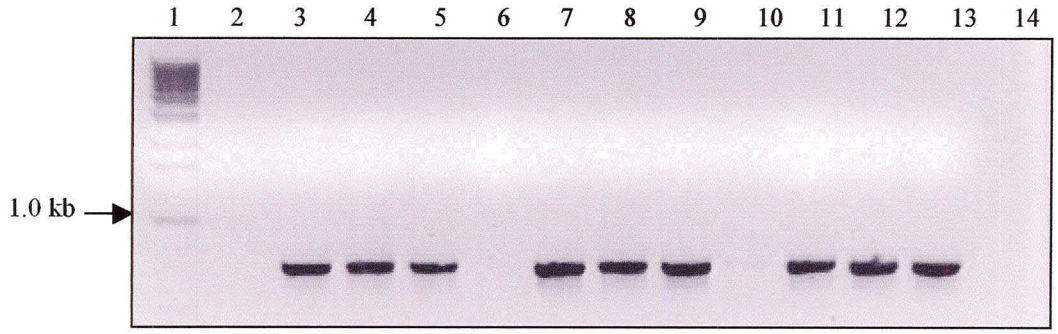


Figure 4.7. Glucocerebrosidase (GBA) and enhanced green fluorescent protein (EGFP) RNA isolated from HeLa cells at 48 hours after co-transfection with small interfering RNAs (siRNA), subjected to RT-PCR using EGFP primers, and electrophoresed in a 1.0% agarose gel. Lanes 2-5 included control siRNA. Lanes 6-9 included EGFP siRNA. Lanes 10-13 included TCP80 siRNA. From left to right (lanes 1-14): lambda ladder, HeLa cells, pEGFP-N1, pGBA1•EGFP, pGBA2•EGFP, HeLa cells, pEGFP-N1, pGBA1•EGFP, pGBA2•EGFP, HeLa cells, pEGFP-N1, pGBA1•EGFP, pGBA2•EGFP, negative (no reverse transcriptase) control.

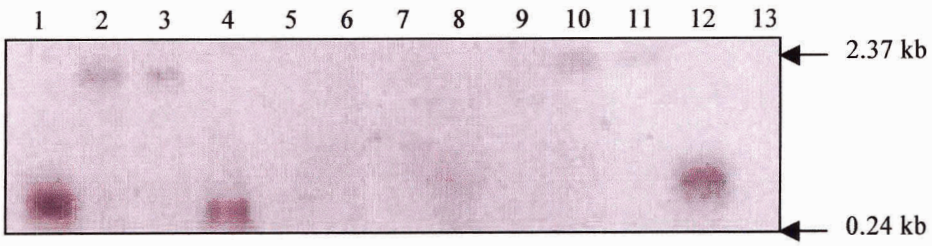


Figure 4.8. Northern blot of glucocerebrosidase (GBA) and enhanced green fluorescent protein (EGFP) RNA isolated from HeLa cells at 48 hours after co-transfection with small interfering RNAs (siRNAs). Lanes 2-5 included TCP80 siRNA. Lanes 6-9 included EGFP siRNA. Lanes 10-13 included control siRNA. From left to right (lanes 1-13): pEGFP-N1 positive control, pGBA2•EGFP, pGBA1•EGFP, pEGFP-N1, HeLa cells, pGBA2•EGFP, pGBA1•EGFP, pEGFP-N1, HeLa cells, pGBA2•EGFP, pGBA1•EGFP, pEGFP-N1, HeLa cells.

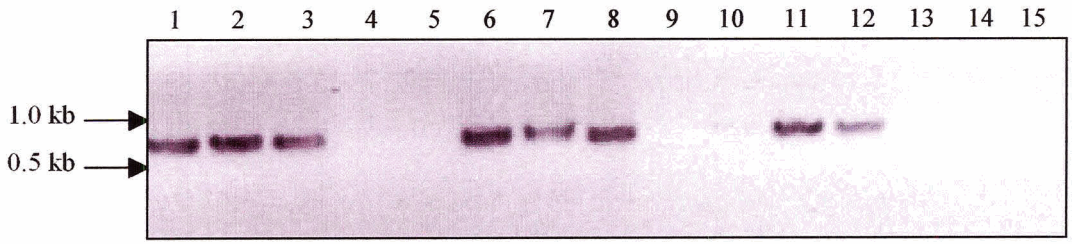


Figure 4.9. Glucocerebrosidase (GBA) and enhanced green fluorescent protein (EGFP) RNA isolated from COS-1 cells at 48 hours after co-transfection with small interfering RNAs (siRNA), subjected to RT-PCR using EGFP primers, and electrophoresed in a 0.7% agarose gel. Lanes 1-5 included control siRNA. Lanes 6-10 included EGFP siRNA. Lanes 11-15 included TCP80 siRNA. From left to right (lanes 1-15): pEGFP-N1, pGBA1•EGFP, pGBA2•EGFP, COS-1 cells, negative (no reverse transcriptase) control, pEGFP-N1, pGBA1•EGFP, pGBA2•EGFP, COS-1 cells, negative control, pEGFP-N1, pGBA1•EGFP, pGBA2•EGFP, COS-1 cells, negative control.

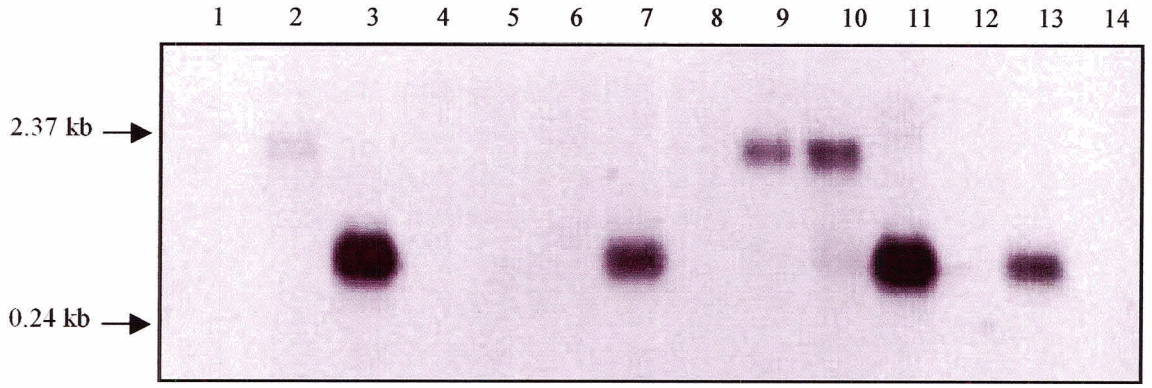


Figure 4.10. Northern blot of glucocerebrosidase (GBA) and enhanced green fluorescent protein (EGFP) RNA isolated from COS-1 cells at 48 hours after co-transfection with small interfering RNAs (siRNAs). Lanes 1-4 included TCP80 siRNA. Lanes 5-8 included EGFP siRNA. Lanes 9-12 included control siRNA. From left to right (lanes 1-14): pGBA2•EGFP, pGBA1•EGFP, pEGFP-N1, COS-1 cells, pGBA2•EGFP, pGBA1•EGFP, pEGFP-N1, COS-1 cells, pGBA2•EGFP, pGBA1•EGFP, pEGFP-N1, COS-1 cells, pEGFP-N1 positive control, COS-1 negative control.

each time point following 24 hours post-transfection in HeLa cells. No fluorescence was observed at any time in control HeLa cells. Fluorescence was also visible 48-72 hours post-transfection in both the pGBA3·EGFP- and pGBA4·EGFP-expressing cultures (Figure 4.11). Consistently greater numbers of fluorescing cells were observed in pGBA4·EGFP-expressing cultures.

4.3.5 Protein Analysis of Co-transfected HeLa and COS-1 Cells

Green fluorescence was regularly observed in pEGFP-N1-expressing cultures at each time point following 24 hours post-co-transfection with either control siRNA or TCP80 siRNA in HeLa and COS-1 cells. Co-transfection of pEGFP-N1 with EGFP siRNA resulted in markedly decreased fluorescence in both cell lines. No fluorescence was observed at any time in control HeLa and COS-1 cells (Figure 4.12; Figure 4.13). With respect to the pGBA1·EGFP- and pGBA2·EGFP-expressing cells, decreased fluorescence was observed when co-transfected with either EGFP or TCP80 siRNA in both cell lines (Figure 4.12; Figure 4.13). However, some differences existed between the HeLa and COS-1 lines. In HeLa cultures, EGFP and TCP80 siRNA seemed to knock down both pGBA1·EGFP and pGBA2·EGFP equally (Figure 4.12). In COS-1 cultures, however, pGBA1·EGFP seemed more affected by EGFP siRNA, whereas pGBA2·EGFP was more affected by TCP80 siRNA (Figure 4.13).

4.4 Discussion

4.4.1 Truncated Plasmid Approach

Despite the use of highly efficient transcriptional promoters and high gene copy

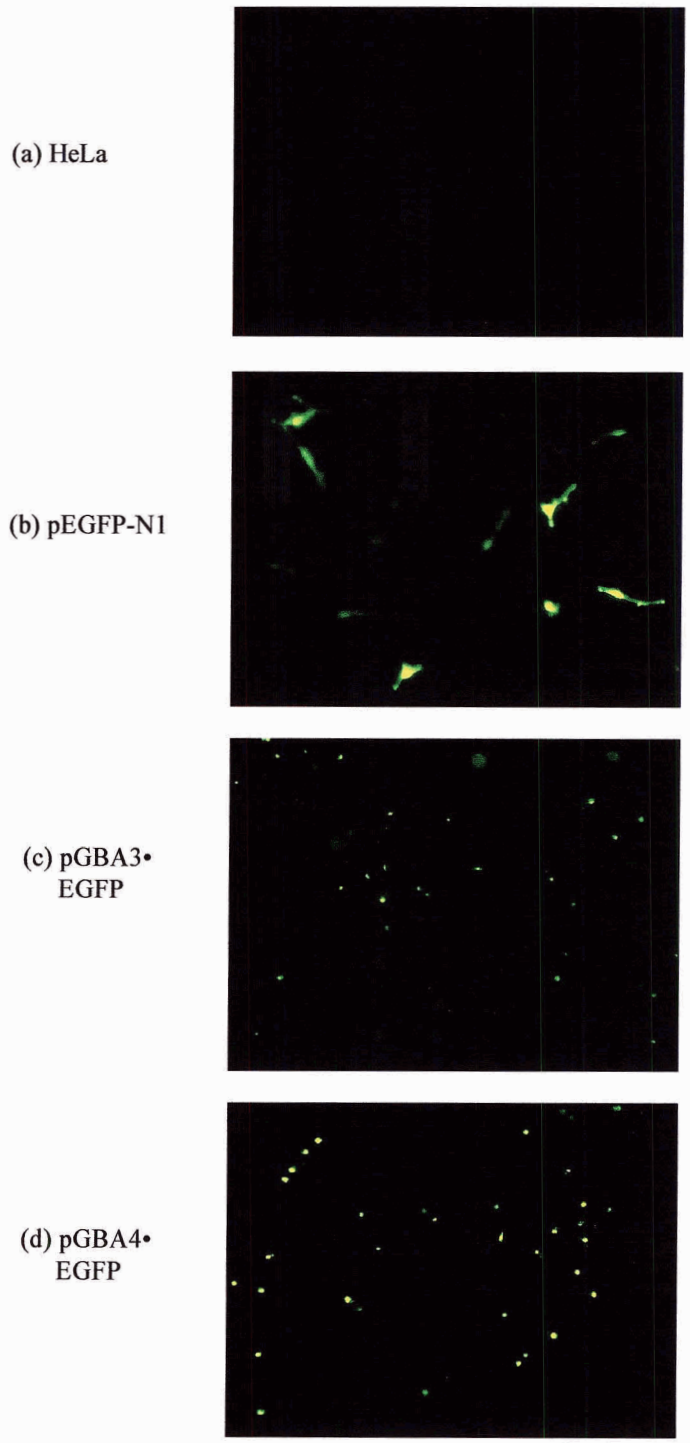


Figure 4.11. Truncated glucocerebrosidase (GBA) and enhanced green fluorescent protein (EGFP) expression in HeLa cells 72 hours post-transfection. Figures (a) and (b) were photographed at 40X magnification, while (c) and (d) were photographed at 20X magnification.

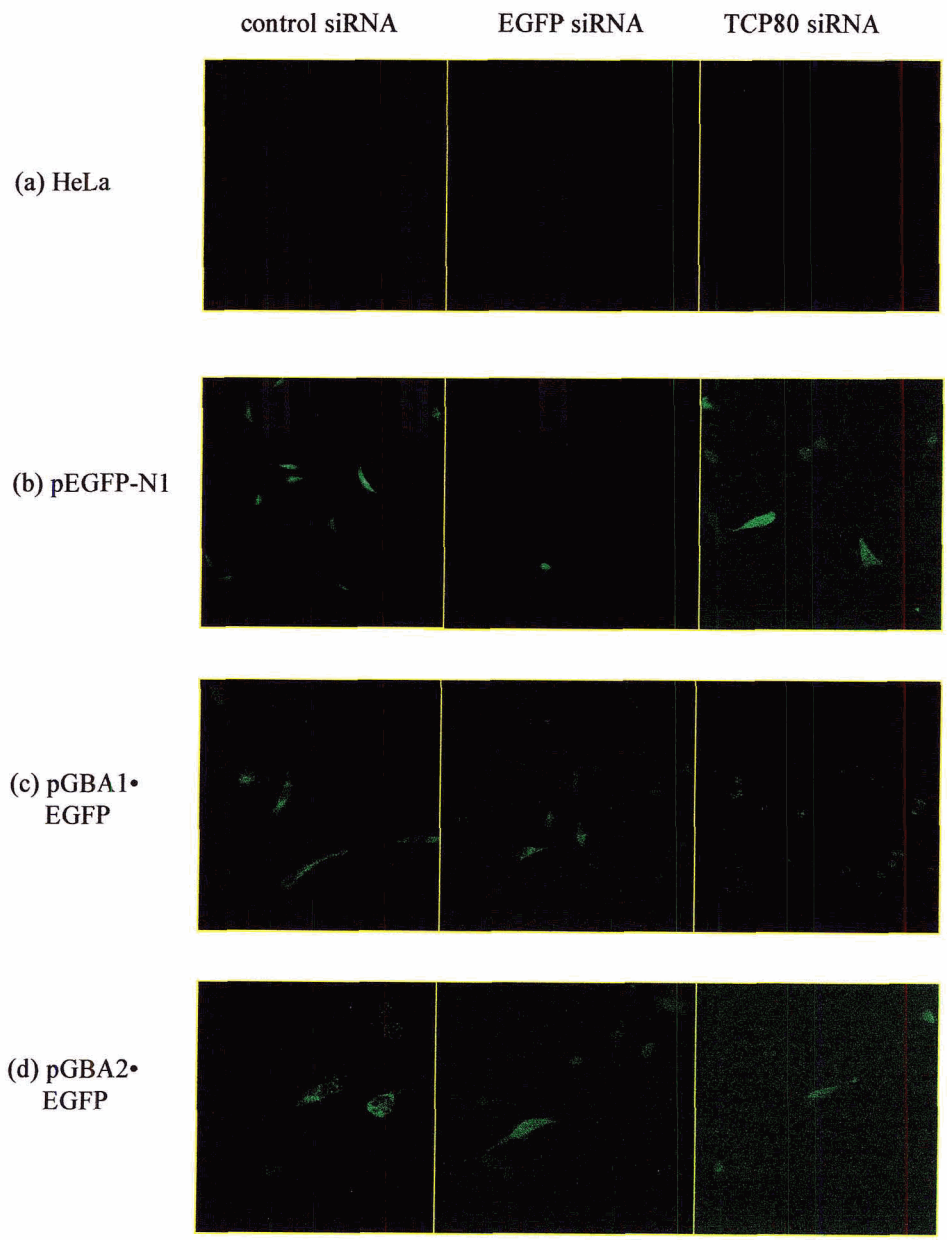


Figure 4.12. Glucocerebrosidase (GBA) and enhanced green fluorescent protein (EGFP) expression in HeLa cells 72 hours following co-transfection with small interfering RNAs (siRNAs).

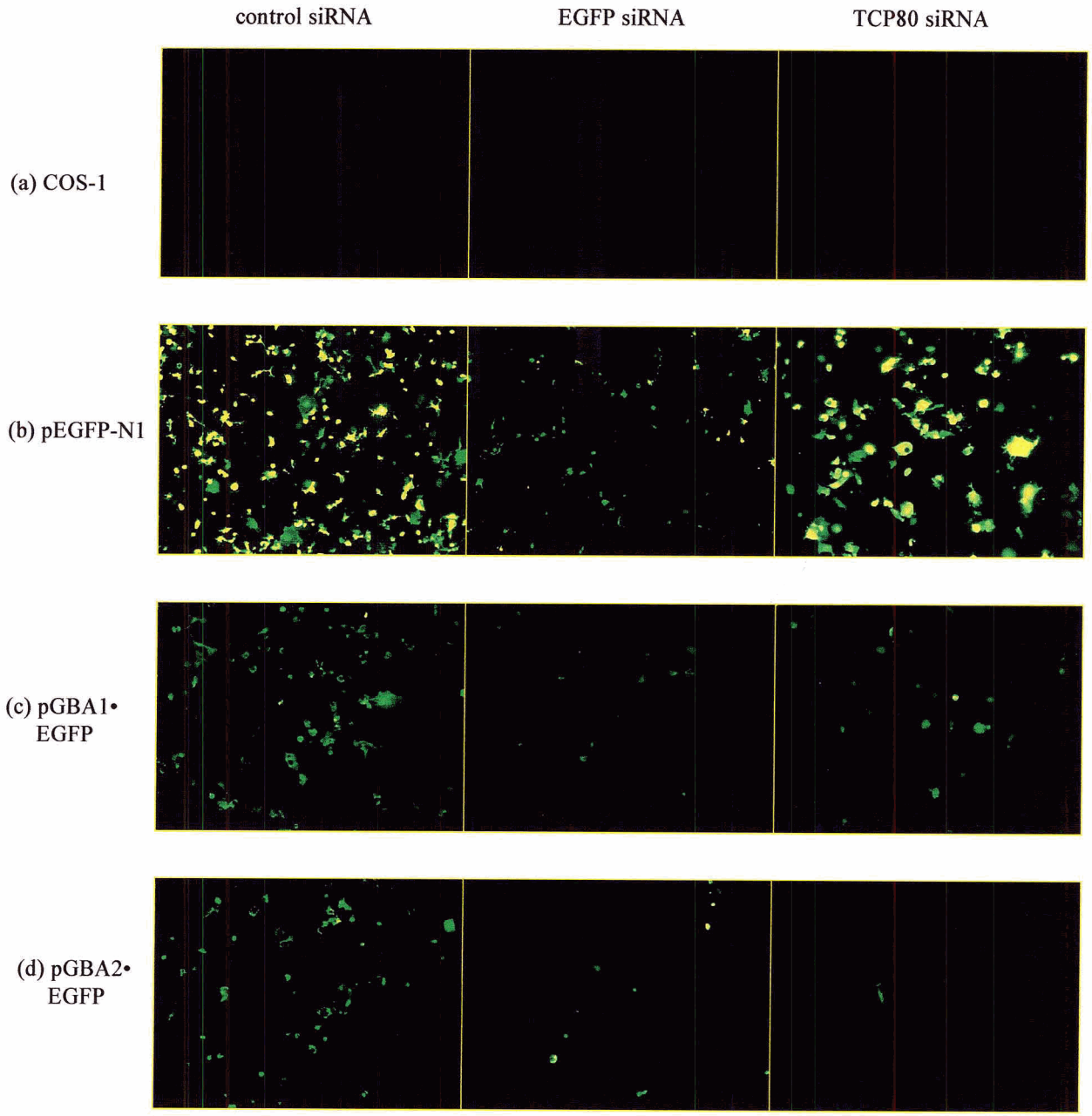


Figure 4.13. Glucocerebrosidase (GBA) and enhanced green fluorescent protein (EGFP) expression in COS-1 cells 72 hours following co-transfection with small interfering RNAs (siRNAs).

strategies, resultant levels of glucocerebrosidase production have remained low in many expression systems (Mistry et. al. 1996, Xu and Grabowski 1998). In mammalian cells, this discordance between transcriptional and translational levels has been proposed to result from inhibition of glucocerebrosidase translation by TCP80 (Xu and Grabowski 1998, Xu et. al. 2000). TCP80 is an mRNA-binding protein with near-identical sequence homology to NF90, MPP4, DRBP76, and ILF3 (Duchange et. al. 2000, Xu et. al. 2000). TCP80 has been shown to specifically bind a 184 nucleotide region of glucocerebrosidase mRNA, preventing subsequent translation initiation (Xu and Grabowski 1999). Therefore, in order to further examine the proposed role of TCP80 in glucocerebrosidase translational inhibition, two approaches were taken: one utilized truncated constructs, the other used siRNAs.

In the first approach, two different EGFP-tagged 500 bp truncated forms of glucocerebrosidase were used to transfect HeLa cells. One form contained the proposed 184 nucleotide TCP80 binding site, while the second contained a sequence downstream of the binding site. Results indicated strong transcription of both truncated forms, as well as of EGFP alone. RT-PCR revealed bands of expected sizes for each culture, with mock-transfected HeLa negative control cultures showing no amplification. Northern blotting results corroborated RT-PCR results. All RNA analyses therefore confirmed the presence of transcript in all expected cases.

Fluorescence was consistently observed in EGFP-expressing cultures with no fluorescence present in control cells. Greater numbers of cells transfected with the vector devoid of the binding site exhibited more fluorescence than those transfected with the other truncated plasmid. This indicates that the 5' end of glucocerebrosidase containing

the proposed binding site may play a greater role in translational regulation than the 3' end. In cells expressing either truncated construct, greater numbers were observed to be fluorescing than in the EGFP-tagged full-length glucocerebrosidase (pGBA1·EGFP)- and EGFP-tagged glucocerebrosidase starting at the second ATG (pGBA2·EGFP)-expressing cultures described in Chapter 3. This therefore suggests that a property of the glucocerebrosidase gene as a whole may also be contributing to the low levels of translation in mammalian cells, rather than only the 184 nucleotide binding region suggested by Xu and Grabowski (1999). Posttranslational modifications have been examined as possible factors influencing glucocerebrosidase expression. Studies, however, indicate that such modifications do not play a significant role in the control of glucocerebrosidase expression and maintenance of steady-state levels (Zhao and Grabowski 2002). Occupancy of the first glycosylation site is essential for enzyme activity, but not for proteolytic stability (Berg-Fussman et. al. 1993). Thus, the complex conditions and sets of factors influencing glucocerebrosidase expression remain to be elucidated.

4.4.2 siRNA Approach

In the second approach, three different siRNAs were employed: one control, one directed against EGFP, and one directed against TCP80. When directly testing for the influence of such siRNAs on TCP80 at the transcriptional level, RT-PCR indicated the presence of transcript in all cultures. Northern blotting, however, showed no cross-reactivity, indicating transcript levels below detection. Thus, any quantitative comparison of different siRNA effects on TCP80 itself was not possible.

When testing for the influence of the three siRNAs on EGFP and EGFP-tagged glucocerebrosidase in HeLa cells, differences were observed. Though RT-PCR results indicated transcript presence in all plasmid-transfected cells, Northern blotting showed no cross-reactivity with any culture that had been co-transfected with EGFP siRNA. This suggests that the EGFP siRNA knocked down transcription of both EGFP and EGFP-tagged constructs. TCP80 siRNA did not appear to have an effect at the transcriptional level. At the translational level, EGFP siRNA co-transfected cultures exhibited decreased numbers of fluorescing cells when compared to control siRNA co-transfected cultures. TCP80 siRNA-containing cultures also showed decreased fluorescence, indicating knockdown at the translational level. This apparent decrease in translation associated with TCP80 siRNA contrasts the expected effect of increased glucocerebrosidase production due to decreased inhibition of translational initiation.

In COS-1 cells, RT-PCR results revealed transcript presence in all plasmid-transfected cells with the exception of pGBA2·EGFP when co-transfected with TCP80 siRNA. Northern blotting showed cross-reactivity with all pEGFP-N1-transfected cultures, with qualitatively less intensity present in the EGFP siRNA co-transfected culture. Cross-reactivity was also visible in both glucocerebrosidase-containing plasmids/control siRNA co-transfected cultures. No bands were observed for glucocerebrosidase-containing plasmids/EGFP siRNA co-transfected cultures, indicating that transcript levels were below detection. A faint band of cross-reactivity was observed in the pGBA1·EGFP/TCP80 siRNA co-transfected culture. Complementing RT-PCR results, no band of cross-reactivity in the pGBA2·EGFP/TCP80 siRNA co-transfected culture was present in the Northern blot, suggesting blocked transcription or levels well

below detection. At the translational level, all plasmid-transfected cultures demonstrated fewer numbers of fluorescing cells when co-transfected with EGFP siRNA, paralleling results previously observed in HeLa cells. Although TCP80 siRNA did not appear to influence pEGFP-N1 translation, levels of translation appeared significantly decreased for pGBA1·EGFP and pGBA2·EGFP. These results also paralleled those seen in HeLa cells. One main difference, however, was that EGFP and TCP80 siRNA seemed to knock down translation of both glucocerebrosidase plasmids equally in HeLa cells, whereas pGBA1·EGFP seemed more affected by EGFP siRNA and pGBA2·EGFP seemed more affected by TCP80 siRNA in COS-1 cells. Additionally, greater transfection efficiencies were consistently observed in COS-1 cells.

The observation that co-transfected HeLa and COS-1 cells did not respond identically is interesting, but not unique. Numerous studies have demonstrated differences in transfection, transcription, translation, and/or processing among various cell lines, and even within a single cell line (Xu and Grabowski 1998, Qiu et. al. 1999, Young et. al. 2000, Kimchi-Sarfaty et. al. 2002, McManus and Sharp 2002). For example, in one study, attempts to increase nuclear export of human immunodeficiency virus type 1 Gag RNA by adding the constitutive transport element (CTE) yielded only a moderate increase in Gag expression by COS cells and an even lower increase in Gag expression in HeLa cells or several murine lines (Qiu et. al. 1999). In another example, introduction of the same siRNAs into NIH/3T3, COS-7, and HeLa cells resulted in knockdown of expression ranging from 9-25-fold (Elbashir et. al. 2001). Numerous explanations have been proposed to account for the variation among cell types, including differences in post-translational processing, folding, co-factor/receptor abundance, and passage number

(Xu and Grabowski 1998, Kimchi-Sarfaty 2002, McManus and Sharp 2002). Whatever the explanation, such observations highlight the usefulness of employing more than one cell type in gene expression studies.

It should be noted that, in general, greater numbers of fluorescing cells were observed for pGBA1·EGFP in comparison with pGBA2·EGFP. This agrees with previous studies of glucocerebrosidase. Sorge et. al. (1987), for example, reported that initiation from the upstream ATG produced two to three times as much protein as did initiation from the downstream ATG. Pasmanik-Chor et. al. (1996) noted a similar phenomenon when measuring *in vitro* glucocerebrosidase activity of both constructs.

Finally, the observation that the introduction of TCP80 siRNA resulted in additional knockdown of glucocerebrosidase expression, rather than increased translation merits further discussion. Originally, TCP80 was reported to inhibit glucocerebrosidase translation (Xu and Grabowski 1998), hence knockdown of TCP80 should increase such translation. The fact that the opposite occurred may be due to a combination of several possibilities. The first is that the siRNA directed against TCP80 did not successfully knock down TCP80, thus failing to decrease its inhibitory effects on glucocerebrosidase expression. Since TCP80 RNA levels were below detection on the Northern blot and since no antibody was available to test expression, this possibility remains to be validated. Because the TCP80 siRNA did appear to have an effect (albeit negative) on glucocerebrosidase expression, it is possible that the siRNA knocked down another gene involved in the regulation of such expression or that TCP80 and its apparent near-homologues play a more broad cellular role than the specific inhibition of glucocerebrosidase translation. Currently, a debate exists whether a number of proteins

with nearly identical sequences are actually the same or are different entities (Duchange et. al. 2000). These proteins include: NF90, ILF3, MPP4, TCP80, and DRB76. All are reported to be double-stranded RNA-binding proteins whose functions have not been fully elucidated. Duchange et. al. recently suggested that all these proteins could be differentially spliced and polyadenylated forms generated from the *ILF3* gene. These forms, therefore, could perform different roles such as the regulation of transcription (as suggested by the presence of NF90 in DNA regulatory complexes; Duchange et. al. 2000), interaction with interferon-induced protein kinase (PKR; Patel et. al. 1999), and phosphorylation during the M phase of the cell cycle (Matsumoto-Taniura et. al. 1996, Duchange et. al. 2000). Hence, the introduction of TCP80 siRNA may have knocked down a similarly sequenced but different protein with a larger role in the regulation of glucocerebrosidase translation or even with a more broad role in general cellular processes. The observation that EGFP, a cytoplasmic protein, did not appear affected by the presence of TCP80 siRNA hints at a possible role of the knocked down gene in the biosynthetic pathway.

4.5 References

Berg-Fussman, A., Grace, M.E., Ioannou, Y., and Grabowski, G.A. (1993). Human Acid β-glucosidase: *N*-glycosylation site occupancy and the effect of glycosylation on enzyme activity. *J. Biol. Chem.* 268: 14861-14866.

Caplen, N.J., Parrish, S., Imani, F., Fire, A., and Morgan, R.A. (2001). Specific inhibition of gene expression by small double stranded RNAs in invertebrate and vertebrate systems. *Proc. Natl. Acad. Sci. USA* 98: 9742-9747.

Chuang, C-F. and Meyerowitz, E.M. (2000). Specific and heritable genetic interference by double-stranded RNA in *Arabidopsis thaliana. Proc. Natl. Acad. Sci. USA* 97: 4985-4990.

Duchange, N., Pidoux, J., Camus, E., and Sauvaget, D. (2000). Alternative splicing in the human interleukin enhancer binding factor 3 *(ILF3)* gene. *Gene* 261: 345-353.

Elbashir, S.M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K., and Tuschl, T. (2001). Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* 411: 494-498.

Elbashir, S.M., Harborth, J., Weber, K., and Tuschl, T. (2002). Analysis of gene function in somatic mammalian cells using small interfering RNAs. *Methods* 26: 199-213.

Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E., and Mello, C.C. (1998). Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans. Nature* 391: 806-811.

Kimchi-Sarfaty, C., Ben-Nun-Shaul, O., Rund, D., Oppenheim, A., and Gottesman, M.M. (2002). *In vitro*-packaged pseudovirions as highly efficient vectors for gene transfer. *Hum. Gene Ther.* 13: 299-310.

Matsumoto-Taniura, N., Pirollet, F., Monroe, R., Gerace, L., and Westerndorf, J.M. (1996). Identification of novel M phase phosphoproteins by expression cloning. *Mol. Biol. Cell* 7: 1455-1469.

McManus, M.T. and Sharp, P.A. (2002). Gene silencing in mammals by small interfering RNAs. *Nat. Rev. Genet.* 3: 737-747.

Mistry, P.K., Wraight, E.P., and Cox, T.M. (1996). Therapeutic delivery of proteins to macrophages: implications for treatment of Gaucher's disease. *Lancet* 348: 1555-1559.

Pasmanik-Chor, M., Elroy-Stein, O., Aerts, H., Agmon, V., Gatt, S., and Horowitz, M. (1996). Overexpression of human glucocerebrosidase containing different-sized leaders. *Biochem. J.* 317: 81-88.

Patel, R.C., Vestal, D.J., Xu, Z., Bandyopadhyay, S., Guo, W., Erme, S.M., Williams, B.R.G., and Sen, G.C. (1999). DRBP76, a double-stranded RNA-binding nuclear protein, is phosphorylated by the interferon-induced protein kinase, PKR. *J. Biol. Chem.* 274: 20432-20437.

Qiu, J-T., Song, R., Dettenhofer, M., Tian, C., August, T., Felber, B.K., Pavlakis, G., and Yu, X-F. (1999). Evaluation of novel human immunodeficiency virus type 1 Gag DNA vaccines for protein expression in mammalian cells and induction of immune responses. *J. Virol.* 73: 9145-9152.

Scherr, M., Morgan, M.A., and Eder, M. (2003). Gene silencing by small interfering RNAs in mammalian cells. *Curr. Med. Chem.* 10: 245-256.

Shi, Y. (2003). Mammalian RNAi for the masses. *Trends Genet.* 19: 9-12.

Sorge, J.A., West, C., Kuhl, W., Treger, L., and Beutler, E. (1987). The human glucocerebrosidase gene has two functional ATG initiator codons. *Am. J. Hum. Genet.* 41: 1016-1024.

Xu, Y-H., Busald, C., and Grabowski, G.A. (2000). Reconstitution of TCP80/NF90 translation inhibition activity in insect cells. *Mol. Genet. Metab.* 70: 106-115.

Xu, Y-H. and Grabowski, G.A. (1998). Translational inefficiency of acid β-glucosidase mRNA in transgenic mammalian cells. *Mol. Genet. Metab.* 64: 87-98.

Xu, Y-H. and Grabowski, G.A. (1999). Molecular cloning and characterization of a translational inhibitory protein that binds to coding sequences of human acid β-glucosidase and other mRNAs. *Mol. Genet. Metab.* 68: 441-454.

Young, P.J., Le, T.T., Thi Man, N., Burghes, A.H.M., and Morris, G.E. (2000). The relationship between SMN, the spinal muscular atrophy protein, and nuclear coiled bodies in differentiated tissues and cultured cells. *Exper. Cell Res.* 256: 365-374.

Zhao, H. and Grabowski, G.A. (2002). Human genome and diseases: review. Gaucher disease: perspectives on a prototype lysosomal storage disorder. *Cell. Mol. Life Sci.* 59: 694-707.

Chapter 5 – Summary of Glucocerebrosidase Expression and Analysis

5.1 Gaucher Disease Background

Gaucher disease, the most common type of lysosomal storage disorder, is characterized by an inherited deficiency of the membrane-associated hydrolase, glucocerebrosidase. Glucocerebrosidase catalyzes the hydrolysis of glucocerebroside to ceramide and glucose, a crucial step in the recycling of membrane sphingolipids. A deficiency in this enzyme results in debilitating symptoms such as enlarged spleen and liver, bone crises, anaemia, and, in some cases, neurological impairment (Beutler and Grabowski 2001). As a single gene enzyme deficiency, Gaucher disease is an ideal candidate for enzyme replacement therapy. Such therapy exists, though the exorbitant cost prevents many from receiving treatment (Grabowski et. al. 1998). Other forms of treatment, such as bone marrow transplant, have also proven successful; however, the high risks associated with such treatment must be balanced with the possible benefits (Hoogerbrugge et. al. 1995). Currently, gene therapy approaches are receiving much attention, though no confirmed successful clinical trials for this form of treatment have been demonstrated to date.

5.2 Heterologous Expression of Glucocerebrosidase-Green Fluorescent Protein Chimerae in *Pichia pastoris*

In an attempt to develop a more cost-effective method of producing glucocerebrosidase, the Choy lab examined expression in *Pichia pastoris*. Unfortunately, low levels of expression were observed. Alterations in growth conditions, plasmids, and

cassette copy numbers did not result in desired levels of expression. To provide further insight into optimizing glucocerebrosidase expression, a new approach was needed. The development of green fluorescent protein (GFP) as a molecular reporter provided such an opportunity to study target protein production through GFP chimerae. Different mutational approaches have optimized GFP expression by altering promoters, codons, splicing patterns, folding, and emission/excitation (Rizzuto et. al. 1996, Tsien 1998). Therefore, two different GFP variants were fused to glucocerebrosidase, serving as *in vivo* markers. Protein analysis indicated that glucocerebrosidase chimerae were not produced at detectable levels. These results corroborate the difficulty of enzyme expression previously noted in the Choy lab, suggesting that *Pichia pastoris* is not currently an appropriate system for glucocerebrosidase production. In contrast, both variants of GFP were successfully produced in the yeast system. Unexpectedly, enhanced green fluorescent protein (EGFP) appeared to be secreted in a greater quantity than red-shifted GFP (RSGFP). Since CAI values indicate that RSGFP should be produced in higher levels than EGFP, the results of this dissertation do not support the codon bias theory, nor do they support suggestions to avoid EGFP expression in yeast (Campbell and Choy 2002).

5.3 Examination of Glucocerebrosidase and Enhanced Green Fluorescent Protein Biosynthesis in HeLa Cells

In order to study glucocerebrosidase biosynthesis and trafficking in a higher eukaryotic system, EGFP-tagged enzyme beginning at either the first or second ATG initiation codon was expressed in HeLa cells. RNA analysis indicated successful

transcription of both forms of glucocerebrosidase, as well as of EGFP alone. These results represent the first confirmation of successful EGFP-tagged glucocerebrosidase at the RNA level. Results from protein analyses, however, were less convincing. Though EGFP was readily visualized, few cells expressing glucocerebrosidase constructs were present. Moreover, no co-localization with organelle markers was evident. A qualitative comparison of RNA versus protein levels suggested an apparent translational inefficiency, supporting evidence from other studies (Xu and Grabowski 1998, Xu et. al. 2000). This translational inefficiency has been proposed to result from the interaction of the glucocerebrosidase transcript with TCP80, an 80,000 MW mRNA-binding translational control protein (Xu and Grabowski 1998, Xu et. al. 2000). TCP80 has been shown to specifically bind a 184 nucleotide region (261 bp - 444 bp) of glucocerebrosidase mRNA, preventing interaction of the mRNA with polysomes and subsequent translation initiation (Xu and Grabowski 1999).

5.4 Examination of Glucocerebrosidase and Enhanced Green Fluorescent Protein Biosynthesis in HeLa and COS-1 Cells Through Truncated Constructs and siRNA Interference

To further examine the proposed role of TCP80 in glucocerebrosidase translational inhibition, two approaches were taken: one utilized truncated constructs, the other employed small interfering RNAs (siRNAs). In the first approach, greater expression was observed from cultures transfected with the EGFP-tagged construct devoid of the proposed TCP80 inhibitory binding site than from the EGFP-tagged construct containing the binding site. This suggests that the proposed binding site does

play a role in controlling glucocerebrosidase translation. However, greater numbers of fluorescing cells were observed in cultures transfected with either truncated construct when compared to cultures transfected with glucocerebrosidase starting at the first or second ATG initiation codon. This therefore indicates a more complex mechanism of translational control than strictly inhibition from the proposed binding site. Posttranslational modifications have been examined as possible factors influencing glucocerebrosidase translation. Though occupancy of the first glycosylation sequon is essential for enzyme activity, posttranslational modifications do not appear to play a role in enzyme stability, expression, or maintenance of steady-state levels (Berg-Fussman et. al. 1993, Zhao and Grabowski 2002).

In the second approach, a siRNA was designed to block TCP80. Co-transfection studies of siRNAs (control-, EGFP-, and TCP80-directed) and plasmids (EGFP, EGFP-tagged full-length glucocerebrosidase, and EGFP-tagged glucocerebrosidase beginning at the second ATG initiation codon) were performed in HeLa and COS-1 cells. All constructs were successfully expressed in both cell types when co-transfected with control siRNA. This indicated appropriate lack of interference by the control siRNA. EGFP siRNA co-transfected HeLa and COS-1 cultures exhibited decreased numbers of fluorescing cells, indicating knockdown of both EGFP and EGFP-tagged glucocerebrosidase plasmids. This suggests that genes of interest can be fused to GFP and then knocked down at specific times with GFP siRNA to examine function and expression with little prior knowledge of the gene of interest. When cells were co-transfected with TCP80 siRNA, glucocerebrosidase-containing constructs exhibited decreased fluorescence. This was in contrast with predicted results, since knockdown of

TCP80 was expected to decrease inhibition of glucocerebrosidase, thus increasing expression. Such results imply a greater overall role for TCP80 than strictly inhibition of glucocerebrosidase translation and/or suggest that reported TCP80-like proteins are distinct entities with diversified roles rather than the same protein (Duchange et. al. 2000).

Overall, some differences were observed between HeLa and COS-1 cells. EGFP and TCP80 siRNA appeared to have similar effects in knocking down translation of both forms of glucocerebrosidase in HeLa cells, whereas full length EGFP-tagged glucocerebrosidase seemed more affected by EGFP siRNA and EGFP-tagged glucocerebrosidase starting at the second ATG seemed more affected by TCP80 siRNA in COS-1 cells. Moreover, greater transfection efficiencies were consistently observed in COS-1 cells. Though differences among cell lines have been noted in previous studies (Xu and Grabowski 1998, Qiu et. al. 1999, Young et. al. 2000, McManus and Sharp 2002), these results highlight the importance of utilizing more than one cell type in gene expression studies.

It should be further noted that, in general, greater numbers of fluorescing cells were observed in cultures transfected with full-length glucocerebrosidase than those transfected with glucocerebrosidase beginning at the second ATG initiation site. This agrees with other studies demonstrating two to three times the amount of protein production with the full-length enzyme (Sorge et. al. 1987, Pasmanik-Chor et. al. 1996).

5.5 Future Directions

This dissertation has addressed the enigmatic issue of glucocerebrosidase

biosynthesis and trafficking in both lower and higher eukaryotic systems. Although the results of this work made direct contributions to the understanding of the complex nature of both glucocerebrosidase and GFP, several promising lines of future research have also been revealed. With respect to glucocerebrosidase synthesis for therapeutic purposes, further development of heterologous protein expression systems for the production of human glycosylated proteins could have a significant impact on the quality of life for Gaucher patients. Closer examination of TCP80 and TCP80-like proteins could provide further clues to glucocerebrosidase translational control and shed light upon possible involvement in other cellular pathways. Finally, the use of GFP for temporal knockdown of fused genes of interest offers an opportunity to study poorly characterized genes and has promising implications in pharmaceutical research.

5.6 References

Berg-Fussman, A., Grace, M.E., Ioannou, Y., and Grabowski, G.A. (1993). Human Acid β-glucosidase: *N*-glycosylation site occupancy and the effect of glycosylation on enzyme activity. *J. Biol. Chem.* 268: 14861-14866.

Beutler, E. and Grabowski, GA. (2001). Gaucher disease. In Scriver, C.R., Beaudet, A.L, Sly, W.S., and Valle, D (eds.), *The Metabolic and Molecular Bases of Inherited Disease* (vol. 3, 8th ed., pp. 3635-3668). New York: McGraw-Hill.

Campbell, T.N. and Choy, F.Y.M. (2002). Expression of two green fluorescent protein variants in citrate-buffered media in *Pichia pastoris*. *Anal. Biochem.* 311: 193-195.

Duchange, N., Pidoux, J., Camus, E., and Sauvaget, D. (2000). Alternative splicing in the human interleukin enhancer binding factor 3 *(ILF3)* gene. *Gene* 261: 345-353.

Grabowski, G.A., Leslie, N., and Wenstrup, R. (1998). Enzyme therapy for Gaucher disease: the first 5 years. *Blood Rev.* 12: 115-133.

Hoogerbrugge, P.M., Brouwer, O.F., Bordigoni, P., Ringden, O., Kapaun, P., Ortega, J.J., O'Meara, A., Cornu, G., Souillet, G., Frappaz, D., Blanche, S., and Fischer, A. (1995). Allogeneic bone marrow transplantation for lysosomal storage diseases. *Lancet* 345: 1398-1402.

McManus, M.T. and Sharp, P.A. (2002). Gene silencing in mammals by small interfering RNAs. *Nat. Rev. Genet.* 3: 737-747.

Pasmanik-Chor, M., Elroy-Stein, O., Aerts, H., Agmon, V., Gatt, S., and Horowitz, M. (1996). Overexpression of human glucocerebrosidase containing different-sized leaders. *Biochem. J.* 317: 81-88.

Qiu, J-T., Song, R., Dettenhofer, M., Tian, C., August, T., Felber, B.K., Pavlakis, G., and Yu, X-F. (1999). Evaluation of novel human immunodeficiency virus type 1 Gag DNA vaccines for protein expression in mammalian cells and induction of immune responses. *J. Virol.* 73: 9145-9152.

Rizzuto, R., Brini, M., De Gorgi, F., Rossi, R., Heim, R., Tsien, R.Y., and Pozzan, T. (1996). Double labelling of subcellular structures with organelle-targeted GFP mutants *in vivo*. *Curr. Biol.* 6: 183-188.

Sorge, J.A., West, C., Kuhl, W., Treger, L., and Beutler, E. (1987). The human glucocerebrosidase gene has two functional ATG initiator codons. *Am. J. Hum. Genet.* 41: 1016-1024.

Tsien, R.Y. (1998). The green fluorescent protein. *Annu. Rev. Biochem.* 67: 509-544.

Xu, Y-H., Busald, C., and Grabowski, G.A. (2000). Reconstitution of TCP80/NF90 translation inhibition activity in insect cells. *Mol. Genet. Metab.* 70: 106-115.

Xu, Y-H. and Grabowski, G.A. (1998). Translational inefficiency of acid β-glucosidase mRNA in transgenic mammalian cells. *Mol. Genet. Metab.* 64: 87-98.

Xu, Y-H. and Grabowski, G.A. (1999). Molecular cloning and characterization of a translational inhibitory protein that binds to coding sequences of human acid β-glucosidase and other mRNAs. *Mol. Genet. Metab.* 68: 441-454.

Young, P.J., Le, T.T., Thi Man, N., Burghes, A.H.M., and Morris, G.E. (2000). The relationship between SMN, the spinal muscular atrophy protein, and nuclear coiled bodies in differentiated tissues and cultured cells. *Exper. Cell Res.* 256: 365-374.

Zhao, H. and Grabowski, G.A. (2002). Human genome and diseases: review. Gaucher disease: perspectives on a prototype lysosomal storage disorder. *Cell. Mol. Life Sci.* 59: 694-707.