Structure-assisted discovery of small molecules for targeting and transport across the blood-brain barrier

by J R Exequiel Timbol Pineda

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Chemistry

Montana State University

Abstract:

The existence of a blood-brain barrier (BBB) prevents more than 95% of therapeutic drugs from entering the brain. Several strategies are being utilized to target drugs to the Central Nervous System (CNS). Among the elegant approaches being pushed forward is the utilization of endogenous BBB transport system present in brain capillary endothelial cells. The transferrin receptor (TfR) is highly expressed in brain microvascular endothelial cells and is know to transcytose this layer of cells to facilitate the uptake the Fe3+-loaded transferrin (Tf). Furthermore, some antibodies against TfR have been shown to piggyback in the process of transporting Tf across the BBB. The structure of TfR has finally been solved in 1999 and this opened up a new possibility that can be explored for BBB drug delivery. A small molecule that specifically binds inside a vestigial active site of TfR can serve as a generic tag for molecules that need to cross the BBB, as long as it does not perturb the physiological function of TfR. In this project, several virtual screening programs were used to scan the Available Chemicals Database and other virtual chemical libraries for lead compounds. The gene for the ectodomain of the human transferrin receptor was cloned into a pFastBac plasmid. Recombinant protein was overexpressed in SF9 cells using a bacculovirus vector. Isothermal titration microcalorimetry was used to detect binding of the top scoring compounds from virtual screening. Hanging drops had been set up and good quality crystals were obtained for X-ray diffraction studies of the protein-ligand complexes.

STRUCTURE-ASSITED DISCOVERY OF SMALL MOLECULES FOR TARGETING AND TRANSPORT ACROSS THE BLOOD-BRAIN BARRIER

by

J.R. Exequiel Timbol Pineda

A thesis submitted in partial fulfillment
of the requirements for the degree

of

Master of Science

in

Chemistry

MONTANA STATE UNIVERSITY
Bozeman, Montana

August 2002

APPROVAL

of a thesis submitted by

J.R. Exequiel Timbol Pineda

This thesis has been read by each member of the thesis committee and has been found to be satisfactory regarding content, English usage, format, citations, bibliographic style, and consistency, and is ready for submission to the College of Graduate Studies.

Dr. C. Martin Lawrence _____ 8/30/2002
(signature) Date

Approved for the Department of Chemistry and Biochemistry

Dr. Paul A. Grieco _____ 8-30-02
(signature) Date

Approved for the College of Graduate Studies

Dr. Bruce McLeod _____ 9-4-02
(signature) Date

STATEMENT OF PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a master's degree at Montana State University, I agree that the Library shall make it available to borrowers under rules of the Library.

If I have indicated my intention to copyright this thesis by including a copyright notice page, copying is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for permission for extended quotation from or reproduction of this thesis in whole or in parts may be granted only by the copyright holder.

Signature Ezequiel T Pineda

Date 8-30-2002

TABLE OF CONTENTS

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

ABSTRACT

The existence of a blood-brain barrier (BBB) prevents more than 95% of therapeutic drugs from entering the brain. Several strategies are being utilized to target drugs to the Central Nervous System (CNS). Among the elegant approaches being pushed forward is the utilization of endogenous BBB transport system present in brain capillary endothelial cells. The transferrin receptor (TfR) is highly expressed in brain microvascular endothelial cells and is know to transcytose this layer of cells to facilitate the uptake the Fe^{3+}-loaded transferrin (Tf). Furthermore, some antibodies against TfR have been shown to piggyback in the process of transporting Tf across the BBB. The structure of TfR has finally been solved in 1999 and this opened up a new possibility that can be explored for BBB drug delivery. A small molecule that specifically binds inside a vestigial active site of TfR can serve as a generic tag for molecules that need to cross the BBB, as long as it does not perturb the physiological function of TfR. In this project, several virtual screening programs were used to scan the Available Chemicals Database and other virtual chemical libraries for lead compounds. The gene for the ectodomain of the human transferrin receptor was cloned into a pFastBac plasmid. Recombinant protein was overexpressed in SF9 cells using a bacculovirus vector. Isothermal titration microcalorimetry was used to detect binding of the top scoring compounds from virtual screening. Hanging drops had been set up and good quality crystals were obtained for X-ray diffraction studies of the protein-ligand complexes.

CHAPTER 1

Introduction

The number of incidents of diseases of the central nervous system dwarfs the combined mortality from cancer and heart disease. In the United States alone, there are an estimated 80 million people that have some disorder of the brain or spinal cord and who require neurotherapeutics. Based on the number of potential beneficiaries, the neuropharmaceutical area is the largest potential growth sector of the pharmaceutical industry. However, the presence of the so-called blood brain barrier (BBB), that excludes >95% of all drugs in the circulation from entering the brain, poses further challenge to that already difficult task of developing drugs. A dictum that became popular in brain-drug discovery is that a molecule must be lipid-soluble and it should have a molecular weight below the 500g/mole threshold to be able to diffuse freely across the BBB.

What is the Blood Brain Barrier?

The BBB, illustrated in Figure 1 (taken from 1) is a physical and metabolic barrier between the CNS and the systemic circulation, which serves to maintain equilibrium and protect the microenvironment of the brain. The anatomical site of the blood-brain barrier is the endothelial lining of the brain microvasculature. Although astrocyte foot processes invest more than 90% of the microvascular basement membrane, the luminal and antiluminal membrane of the endothelial cells form the main diffusion barrier for solutes crossing from the circulation to brain interstitial fluid. This thin membranous structure,

separated by about 0.3 μm of cytosol, significantly impedes the entry from the blood to brain of virtually all molecules, except those that are small and lipid-soluble. Brain capillaries stretch a length of approximately 400 miles, covering a surface area of approximately $20m^2$. The brain capillaries are approximately $40\mu m$ apart, and it takes around one second for a drug to diffuse $40\mu m$. This allows for almost instantaneous solute equilibration throughout the brain interstitial space once the endothelial barrier is overcome (1-2).

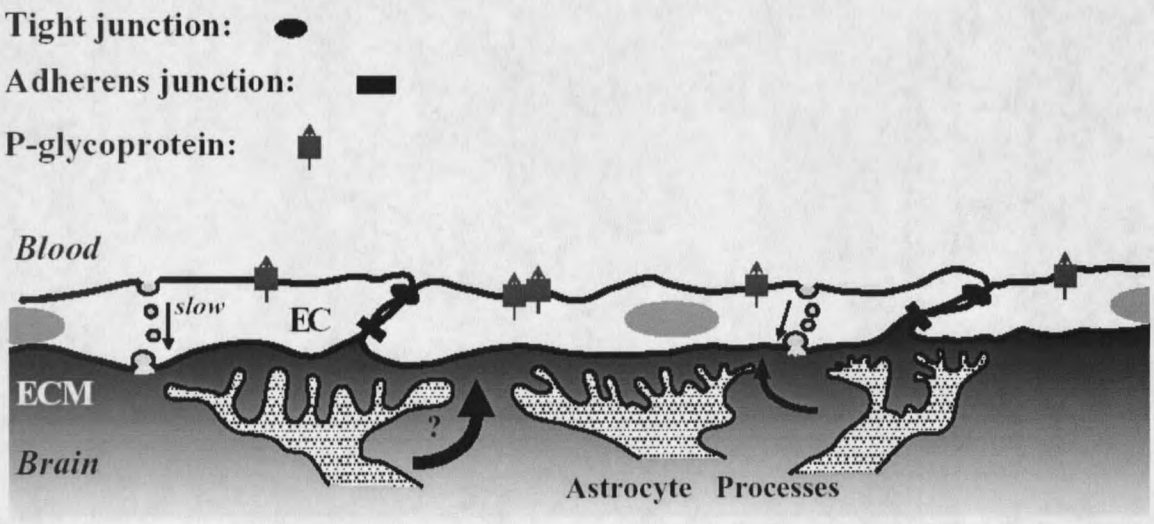

Figure 1. The anatomy of the Blood Brain Barrier

Endothelial cells in the brain differ from those in other peripheral organs in fundamental ways. Fewer endocytic vesicles are detectable by ultrastructural studies, hinting that there is reduced transcellular flux of free solute. Also, the cells are coupled by tight junctions, creating a rate-limiting barrier to paracellular diffusion of solutes between the endothelial cells. The tight junctions are the most apical elements of the junctional complex, which includes both tight and adherens junctions. Their presence is the predominant factor that results in high transendothelial electrical resistance (1500 to

2000 Ω*cm^2) of brain capillary endothelial membrane, a magnitude similar to that of epithelial membranes (1,3).

The high metabolic activity of cells in the central nervous system requires a constant supply of nutrients. There are sets of small and large molecules that can enter the brain via active transport. Membrane transporting proteins for glucose and certain amino acids are present in relatively high concentrations on brain capillaries. Systems capable of transporting macromolecules into the brain are also known, with some of them thought to be receptor-mediated. The best known of these is the transferrin receptor. Another important transporter present at relatively high concentration on brain capillaries is P-glycoprotein. It works in an opposite direction to those previously described in that it transports back into the blood a variety of lipophilic molecules that penetrate the brain or enter the endothelial cells. Circulating drugs and potential toxins concentrate at a higher-than-normal level in the brain of P-glycoprotein knock-out mice, indicating that this membrane protein is a functional component of the barrier (2).

Current Status of Neuropharmaceuticals

At present, the only diseases of the central nervous system that are treatable by small molecule drug therapy are the following:

- Obsessive-compulsive disorders
- Depressions
- Schizophrenia
- Epilepsy
- And Chronic pain

On the other hand, medicines for the following CNS patient groups are scant:

- Alzheimer's disease

- Brain and spinal cord trauma
- Huntington's disease and other neurodegenerative disorders
- Brain cancer
- Stroke
- AIDS
- Infection in the brain
- Genetic diseases that lead to metal retardation and premature death
- Ataxias (Inability to coordinate voluntary muscle movements), etc.

Clearly, only a few diseases of the central nervous system are amenable to small molecule drug therapy. This is, in part, because many promising small molecules are not able to cross the blood-brain barrier (BBB). But more than 98% of small molecules do not cross the blood brain barrier because they fail to satisfy both criteria of lipid-solubility and a molecular weight less than 500g/mole (4).

Trial-and-error methods employed in traditional CNS drug discovery invariably selected drugs that had appropriate drug-receptor interactions and transport properties. In contrast, modern receptor-based high throughout screening methods screen large collections of compounds for leads. But most drug leads that come out of receptor-based high throughput screening (HTS) programs lack the two criteria necessary for BBB transport. It's almost inconceivable that CNS drug discovery has evolved in the absence of a parallel maturation of an effective CNS drug delivery strategy. Indeed, knowing the presence of the BBB, it is quite odd that >99% of worldwide CNS drug development is devoted to CNS drug discovery with <1% of the effort devoted to CNS drug delivery (). Without an effective BBB drug delivery strategy, a HTS program for CNS drug development is destined for termination if it fails to identify compounds that satisfy the traditional standards for BBB permeability.

Ways of Solving the BBB Drug Delivery Problem

The currently proposed solutions to the BBB drug delivery problem can be categorized as (a) invasive or (b) non-invasive strategies. Invasive strategies attempt to either circumvent the BBB by neurosurgical approaches or induce transient disruption of the BBB by osmotic stress. Non-invasive approaches alter the molecular properties of the drug so that it will cross the BBB, with either pharmacologic or physiological strategies.

Invasive Strategies

Intra-cerebroventricular (ICV) infusion or intracerebral implants are the material science solution to the BBB problem in brain drug delivery. Controlled release of drugs in the brain is achieved by implanting a degradable polymer that encapsulates the drug or genetically engineered cells that are designed to secrete the required factor. With ICV, the drug is directly infused into the ventricular compartment with the goal of distributing the drug, mainly by diffusion, throughout the brain parenchyma. Sole reliance on diffusion from the local depot site is the principal problem of both ICV infusion and intracerebral implants because diffusion is a poor mode of drug delivery to the brain. In the case of ICV infusion, most of the drug is rapidly exported from the ventricles to the bloodstream. The drug that does enter the brain is confined only to the ipsilateral ependymal surface of the brain, because of the logarithmic fall in brain-drug concentration relative to the distance (mm) from the surface of the brain. Similarly, the effective diffusion distance of the drug from an intra-cerebral implant is only 1-2 mm. The other huge problem is the cost of the procedure, every ICV implant costs about

US$15,000, and there is no guarantee that implantation to a single locus will result to a complete cure (6).

Another strategy, that is also quite invasive, is transient BBB disruption by injecting a hypertonic solution, i.e. 1.7 M mannitol, or a vasoactive agent, i.e. bradykinin, directly into the carotid artery. The BBB cells act as osmometers, with the cell volume changing inversely as the osmolality of the environment. Cells shrink in a hypertonic environment as the intracellular water moves out. The cells of reduced volume must exert a tension at the tight junctions, thus creating openings, which temporarily modify the permeability of the BBB (7). However, this transient disruption also allows the entry of other molecules that are normally excluded by the healthy BBB, which could potentially cause more harm to the already compromised health of the patient.

Non-invasive Strategies

Drug Lipidization. A pharmacologic approach involve the 'lipidization' of drugs, whereby water-soluble molecules are conjugated to lipid carriers, such as free fatty acids, or replacement of highly polar functional groups with suitable bioisosteres. But although lipidization increases barrier permeability, it also increases drug uptake into the peripheral tissues.

Utilize Endogenous BBB-Transport Systems. An alternative to the invasive and pharmacologic approaches outlined above that is not dependent on lipid solubility and a molecular weight threshold is to reformulate the drug such that it can access the endogenous transport systems that exist in brain capillary endothelial membranes. Three

different classes of endogenous transport systems exist within the BBB. These are: (a) carrier-mediated transport systems (CMT), (b) receptor-mediated transcytosis (RMT) systems, and (c) active efflux transporters (AETs). Carrier-mediated transport systems include the glucose and amino acid carriers that allow bi-directional movement of small molecule nutrients and vitamins between the blood and the brain. The timescale of transport through carrier-mediated transport systems occur on the order of milliseconds. Among the receptor-mediated transcytosis systems are the BBB insulin receptor and the transferrin receptor. These systems mediate the bi-directional movement of large molecules between the blood and the brain and the process is completed within minutes. Active efflux transporters, such as P-glycoprotein, mediate the transport of small molecules from the brain to the blood (8).

The strict steric requirements of most CMT systems make them less attractive for use as delivery targets because a drug conjugated to an endogenous substrate will probably be exclude from the binding sites. Also, most drugs cannot be structurally altered to take on a molecular structure that mimics the endogenous nutrient, and thus most drugs will not have access to the carrier-mediated transport systems within the BBB. An alternative approach is to use the receptor-mediated transcytosis systems within the brain capillary endothelial plasma membrane. RMT is comprised of three steps: (a) receptor-mediated endocytosis at the luminal membrane of the capillary endothelial cells; (b) movement through the 300 nm of endothelial cytoplasm; and (c) exocytosis across the abluminal endothelial membrane into the brain interstitial fluid.

Transferrin Receptors in the BBB

It was demonstrated that monoclonal antibodies against rat and human transferrin receptors (TfR) preferentially label blood capillaries within the brain. Furthermore, in rats, labeling occurs after injection of antibody into the blood, implying that the receptors are accessible at the endothelial surface. From these results, it was suggested that the receptors might be expressed on these cells to allow transport of transferrin into brain tissues (9).

Recent results from a BBB genomics initiative further highlighted the abundance of TfR in rat brain capillary endothelial cells. The Northern blot analysis shown in Figure 2 (taken from 10) demonstrates expression of the 6.6kb transcripts in isolated rat brain capillaries. The mRNA for the full-length rat TfR contains 3413 bases. A recent study have also identified a second form of the transferrin receptor encoded by mRNA of 2.9 and 2.5 kb and is specific for liver (10).

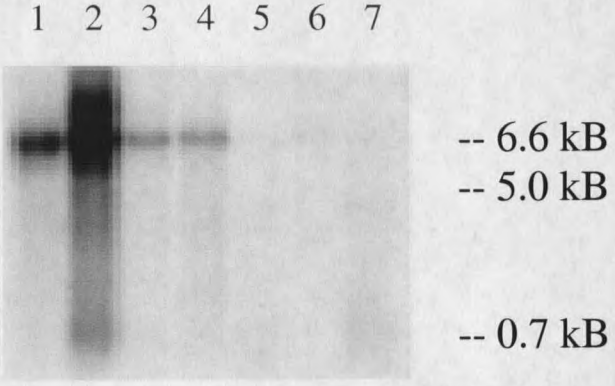

Figure 2. Northern Blot of samples from different tissues. Lanes; (1) C6 rat glioma cells, (2) rat brain capillaries, (3) total rat brain, (4) rat heart, (5) rat kidney, (6) rat lung, and (7) rat liver, respectively

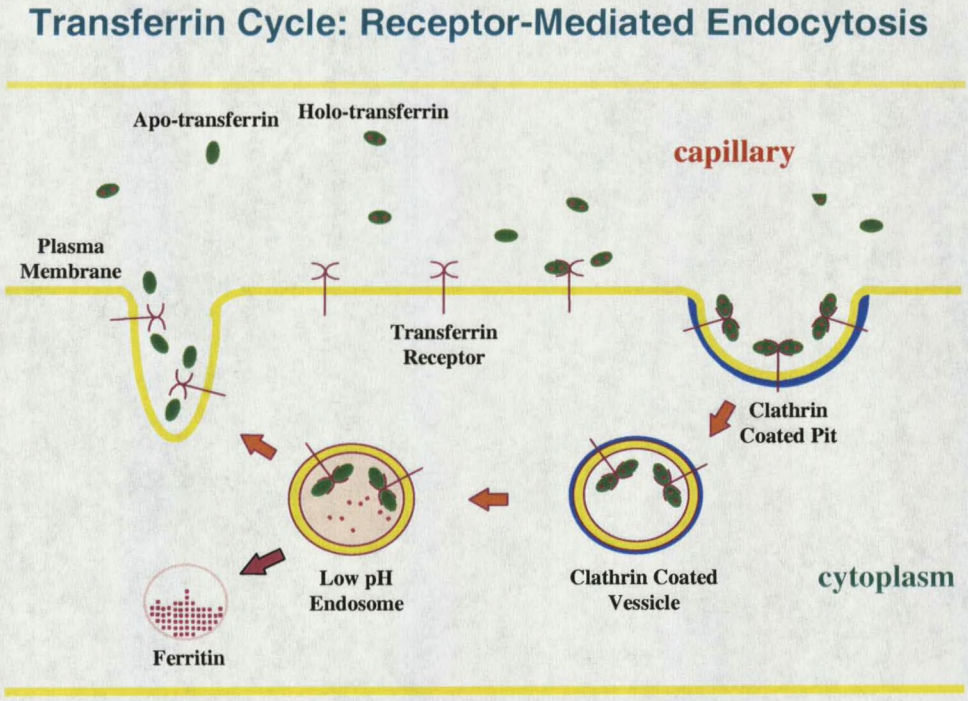

Figure 3. Receptor-mediated uptake of diferric transferrin

The mechanism of iron uptake of most vertebrate cells that require iron is depicted above. Extracellular diferric transferrin binds to the membrane-bound transferrin receptor. The complex is internalized via receptor-mediated endocytosis into an endosome. Acidification of the endosome causes iron to dissociate from transferrin, but apo-Tf remains bound to the receptor. The complex returns to the surface where apo-Tf is released and the receptor becomes free to bind another molecule of diferric transferrin from the serum (11).

OX26 Monoclonal Antibody-Mediated Delivery of Peptides Across the BBB

In one of the initial demonstrations of RMT of Tf across the BBB, radiolabeled Tf was perfused into rat brain. At first, the radiolabel accumulated in the endothelial cells of the BBB followed by a decrease of this intra-endothelial radioactivity associated with a concomitant increase in the nonvascular parts of the brain (12). Years later, an independent study by Skarlatos, et al validated these results. In these later experiments the transport of rat holo-transferrin was evaluated by in situ brain perfusion with the radiolabeled protein. Unlike before, the brain was postperfused with saline prior to decapitation of the rat. They again demonstrated a rapid and extensive distribution of the blood-borne Tf throughout the brain parenchyma. Furthermore the presence of competing plasma Tf in the perfusate reduced the amount radiolabeled Tf detected in the brain (13).

Broadwell, et al, carried out an example of an attempt to target a large biomolecule to the brain. They conjugated diferric-Tf to horseradish peroxidase, which by itself does not cross the BBB. The diferric-Tf-HRP conjugate labeled BBB vessels throughout the CNS without discernible disruption of the BBB or extravasation of the blood-borne probes into the brain parenchyma. After less than one hour, peroxidase activity was detected in brain compartments distal to the endothelial space. (14) These examples clearly illustrate that diferric-Tf is delivered to the brain tissues via receptor mediated transcytosis. The ensuing model is illustrated below.

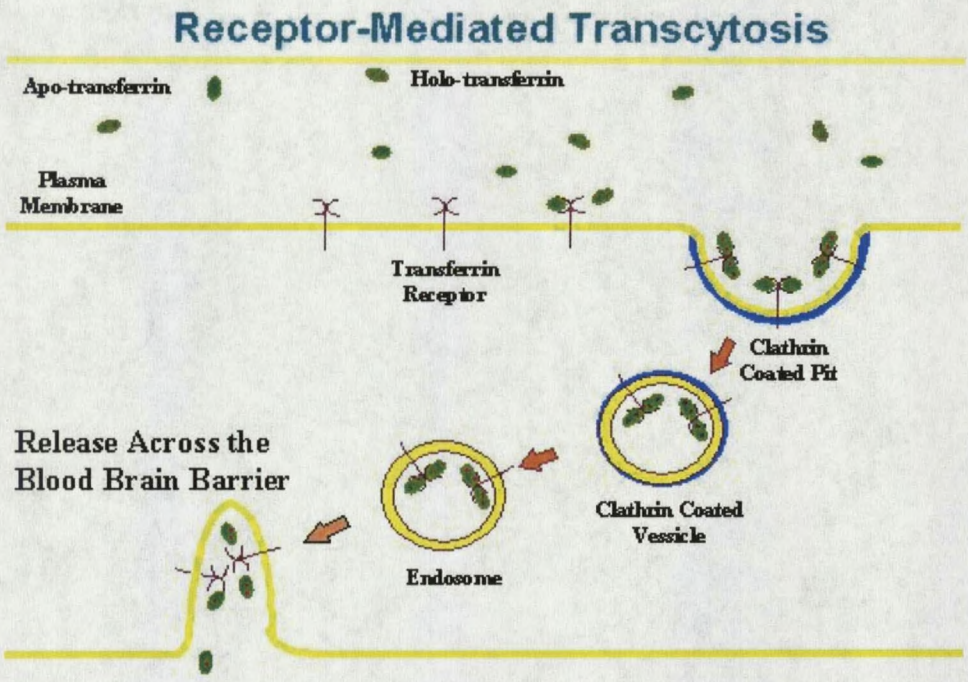

Figure 4. Receptor-mediated transport of diferric transferrin across the BBB endothelial layer

Although diferric-Tf was shown to transcytose better than the OX26 monoclonal antibody against rat TfR, it is still unwise to use it to deliver therapeutic molecules across the BBB. The main impediment to its use is the high level of endogenous diferric-Tf in the serum. A Tf-drug conjugate will have to compete with this pool to gain access to the RMT system in the BBB. To target molecules to BBB TfR the OX26 monoclonal antibody, or its Fab, is still the wiser choice. OX26 binds to an exofacial epitope on rat TfR that is removed from the Tf binding site. Therefore, in rat models of CNS diseases, conjugating this antibody to a therapeutic molecule was shown to be a promising approach to BBB drug delivery with minimal perturbation on the normal receptor function. In fact, peptide-based neuropharmaceuticals conjugated to such monoclonal

antibodies against unique cell surface receptors in the human BBB are subjects of intense studies and are very well documented (15)

Crystal Structure of the Ectodomain of Human TfR

Lawrence, et al in 1999, solved the crystal structure of the extracellular domain of human TfR, corresponding to residues 121 to 760. This fragment is equivalent to that released when TfR-containing membranes are treated with trypsin. The soluble receptor is a dimer and it binds two transferrin molecules. $TfR^{121-760}$ is 28% identical at the amino acid level to membrane glutamate carboxypeptidase II (mGCP) whose substrate, N-acetyl-α-L-aspartyl-L-glutamate, is the most prevalent mammalian neuropeptide (16-17). TfR, however, lacks peptidase activity presumably because it lacks three of the supposed Zinc ligands in the predicted protease-like domain.

The TfR monomer has three distinct domains, and the dimer assumes a butterfly-like shape, as depicted in the Figure 5.

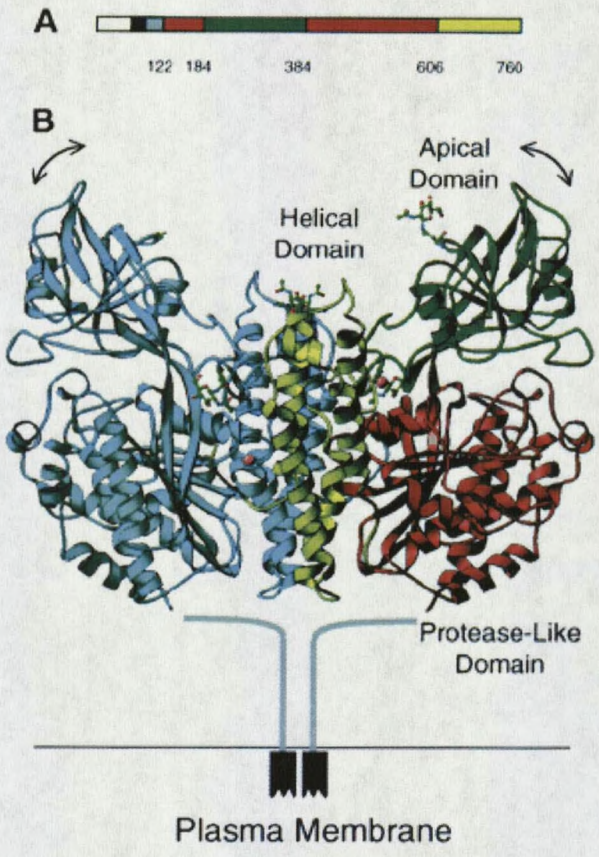

Figure 5. Structure of the ectodomain of human transferrin receptor solved by x-ray crystallography

The domain boundaries in one monomer are color-coded in for clarity, with the protease-like domain in read, helical domain in yellow and the apical domain in green. The protease-like domain has a fold that is closely related to that of carboxy- and aminopeptidase. Its central, seven-stranded, mixed β sheet is flanked by α helices. Eight β strands are present in carboxypeptidase itself, but in TfR the polypeptide chain traces a path away from the outside edge of the β sheet forming an extended loop (see Figure 6).

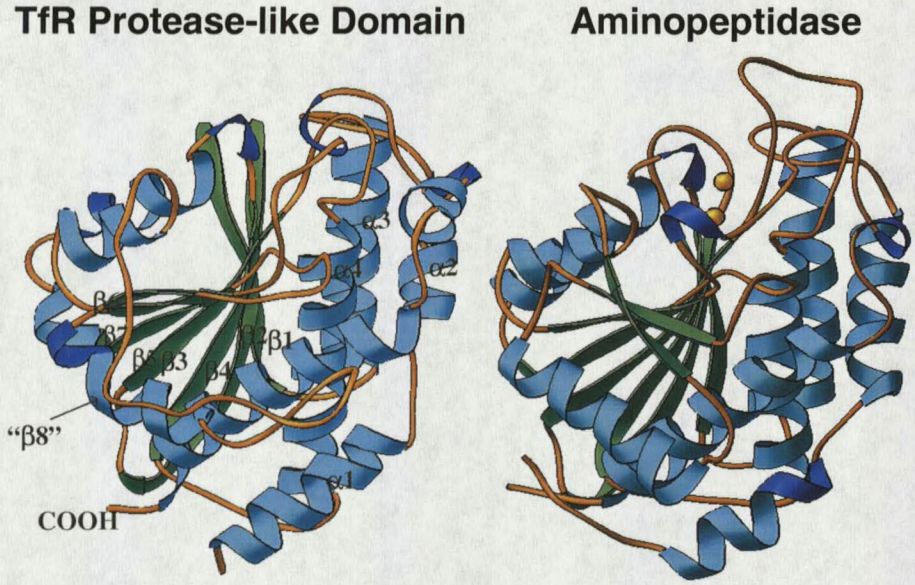

Figure 6. Structural comparison of the protease-like domain of human TfR and an aminopeptidase

The apical domain resembles a β sandwich in which the two sheets are splayed apart, and with a helix running along the open edge. The helical domain is responsible for dimerization. The helical domain of one monomer interacts with its counterpart across the molecular twofold axis, contacting each of the three domains. A significant sequence identity between corresponding domains in TfR and mGCP is revealed by a structure-based alignment. The protease-like, apical, and helical domains of TfR share 30.3%, 30.2%, and 24% identity, respectively, with those of mGCP. The apical domain covers the catalytic site, but the position of the β8 loop allows access through an interdomain channel (11). The location of this channel that leads to the vestigial active site in TfR is more obvious in the surface rendered image in Figure 7. Together with the sequence

similarity, this structural feature suggests that TfR may have evolved from a cell-surface protease, similar to mGCP.

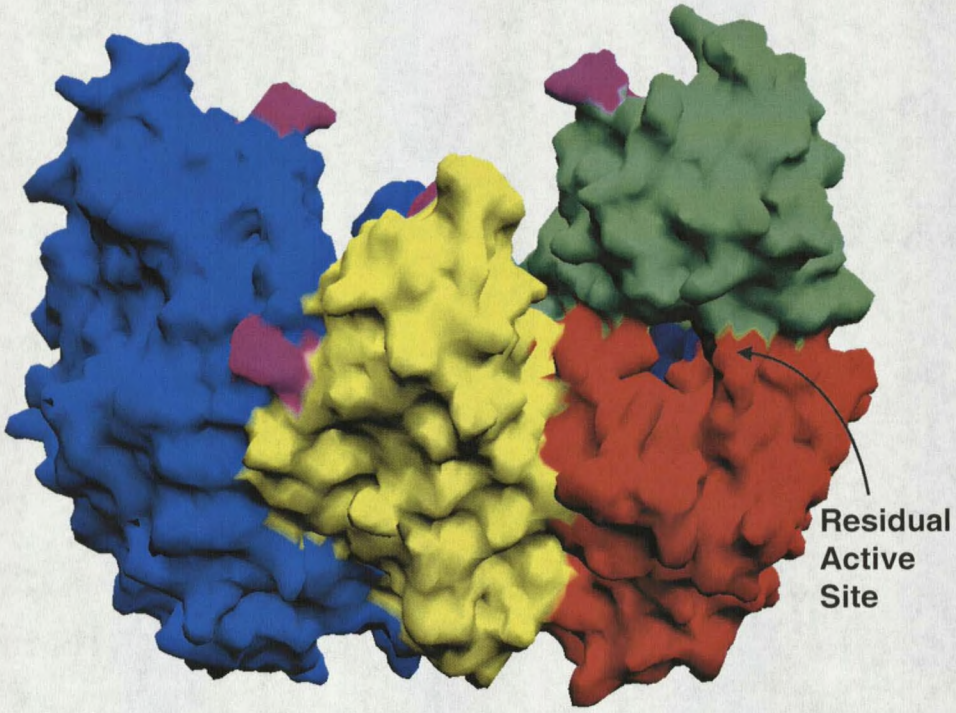

Figure 7. Surface rendered image of the ectodomain of human TfR

It is not clear if this vestigial active site has any functional role. However, its discovery has opened up a new possibility that can be explored for BBB drug delivery: a small molecule that specifically binds inside this vestigial "active site" pocket could serve as a molecular stamp for delivering drugs across the blood-brain barrier, provided that the chimeric molecule does not significantly perturb the physiological role of TfR.

Using the structural features of this pocket as constraints for virtual screening, the program DOCK (Version 4.0) was used to screen libraries of commercially available databases namely: Available Chemicals Database, Tripos and Chembridge. DOCK was

quick but not as good a predictor as more CPU intensive programs, such a FLO. However, the more CPU intensive programs are impractical, with our current computational resources, for screening such libraries that contain about 300,000 compounds. A good compromise is to screen, or filter, the entire database with DOCK. Then the top ranking compounds of approximately 2,500 molecules were evaluated with the more CPU intensive program FLO, using the DOCK-generated conformations as input for energy minimizations in the context of the vestigial active site. The ligands were allowed to move freely while the active site residues were restrained during the minimization. After a detailed examination of the calculated total binding energies and the corresponding component energies, together with visual inspection of the energy minimized conformation of the active site-ligand complexes, we selected a subset of 27 compounds for further in vitro screening.

While these virtual experiments were underway, we overexpressed the ectodomain of human TfR in a lytic insect cell/baculovirus expression system. Some of the protein was used to set up hanging drops for crystallization. The bulk of the isolated protein was used to assay for binding of the selected compounds using isothermal titration calorimerty. Ligands that were found to bind were diffused into TfR crystals for determination of the structure of the complex by x-ray diffraction. It is hoped that with the structure of the protein-ligand complex, we will be able to: (1) determine if the ligands bind to the site where the docking experiments were directed, (2) identify TfR residues that interact directly with the ligand, as well as ligand functional groups that contact the receptor and (3) determine if there are any significant movements on the receptor that are promote binding (induced-fit). These informations will allow us to

propose modifications in the ligand structures that will improve their affinity for TfR, and therefore start the cycle of design and optimization of the leads.

CHAPTER 2

Design of Transfer Vector

Bacterial expression systems are the most established and probably the oldest protein expression systems utilized, however they have an inherent limitation in that they do not have the necessary machinery to process proteins that require post-translational modifications to achieve their functional status. This is true for most eukaryotic proteins expressed on the cell surface. As a result, alternative protein production systems that more closely mimic the complex post-translational modification processes of mammalian cells are being sought for production of complex proteins (18). Insect cells perform post-translational modifications, such as glycosylation, phosphorylation, palmitoylation, myristylation and addition of glycosyl-phosphatidylinositol anchors (19). Glycosylation is the most extensive and necessary modification as it is important in secretion, antigenicity and clearance of glycoproteins (20).

Baculovirus expression vectors provide a versatile and reliable system for the production of recombinant protein in insect cells. One of the most widely used vectors is the *Autographa californica* multiple capsid nucleopolyhedrovirus (AcMNPV), a prototype member of *Baculoviridae*. Several cell lines can be used to propagate this virus; among them are *Estigmene acrea*, *Mamestra brassicae*, *Trichoplusia ni* (Tn) and *Spodoptera frugiperda* (Sf) (21). Previous work have been carried out to examine the extent of N-glycosylation of secreted placental alkaline phosphatase produced by recombinant baculovirus-infected *Trichoplusia ni* and *Spodoptera frugiperda* cell lines.

For this particular protein, it was concluded that *Spodoptera frugiperda* cells produced more fucosylated oligosaccharides than either of the *Trichoplusia ni* cell lines (22).

The early and late genes in the life cycle of AcMNPV are largely concerned with the production of virus particles, which bud from the cell to spread infection to new cells. In contrast, the very late genes, those encoding polyhedrin and p10, are required for the production of occlusion bodies that contain virus particles, in the nucleus of the host cell. Both of these genes are under the control of strong promoters, but can be deleted from the virus genome without affecting the production of infectious virus particles. This is exploited by inserting foreign coding genes in lieu of the polyhedrin and p10 sequences to derive expression vectors (21).

AcMNPV has a large (130-kb), circular, double-stranded DNA genome with multiple recognition sites for many restriction endonucleases. Therefore, it is not wise to construct the recombinant bacmid by traditional "cut and paste" approach. A way around this is to clone the gene of interest in a transfer vector containing a baculovirus promoter flanked by baculovirus DNA derived from a nonessential locus, for example the polyhedrin gene. The gene of interest is inserted into the genome of the parent virus by homologous recombination after transfection in insect cells. This strategy however, requires successive plaque purification to ensure isolation of the pure clone of recombinant virus. A quicker strategy uses site-specific transposition with Tn7 to insert the heterologous gene into bacmid DNA propagated in E. coli. This technique still requires cloning the desired gene into a transfer vector, but the transposition event is carried out in bacterial host cells. Antibiotic selection and blue/white screening identified colonies that contain the recombinant bacmid, because the transposition results in

disruption of the lacZα gene. High molecular weight DNA was prepared from selected clones and this DNA was used to transfect insect cells. Having prepared the DNA from a pure clone, this strategy avoids successive plaque purification to obtain recombinant virus (23).

Design of Baculovirus Transfer Vectors

TfR is a type II integral membrane protein. The ectodomain, composed of residues 121 to 760, contains an intramolecular disulfide bond between Cys^{556} and Cys^{558} (11). Previous structural work utilized protein overexpressed in CHO cells. These cells were expensive to maintain, both in terms of time and money. We wanted to know if we could reproduce the clone produced in CHO cells as closely as possible in a different expression system. We decided to express the protein in a lytic baculovirus/insect cell expression system. The pFASTBAC1 plasmid of the Bac-to-Bac baculovirus expression system was available in the laboratory. However, this transfer vector does not contain a secretion signal sequence and a 6x His tag. A plasmid for a different system, pAcGP67TfR of the BaculoGold system, that contains the secretion signal, 6x His tag, Factor Xa cleavage site and TfR residues 121-760, respectively, was obtained as a generous gift from Tony Gianentti and Prof. Pamela Bjorkman (CalTech). This construct contains a BamH1 endonuclease site between the secretion signal and the 6x His tag. pFASTBAC1 has a BamH1 site in its multiple cloning site.

To avoid having multiple BamH1 cleavage sites in our final construct, we decided to delete this site in pFASTBAC1. 10 μg of the plasmid was digested with 20 units of BamH1. The linearized vector was purified by agarose gel electrophoresis, sticky ends repaired by treating with 10 units of Klenow to generate blunt ends, heated at 60°C to inactivate Klenow and ligated overnight using T4 DNA ligase. MC1061 competent cells were transformed with this ligated DNA and selection was done by streaking several dilutions of the transformed cells on gentamycin-containing agarose plates. A single colony was picked for miniprep and the isolated plasmid was digested with several endonucleases that recognize the multiple cloning site. The difference in hydrodynamic behavior of the linearized as opposed to the circular plasmid DNA indicated the absence of a BamH1 site in the purified plasmid. We refer to this plasmid as pFBΔBH1.

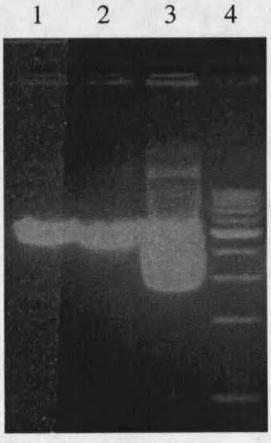

Figure 8. Digestion of pFBΔBH1 with (1) Not1, (2) EcoR1 and (3) BamH1. Lane 4 shows a 1kb ladder

Having pFBΔBH1, we are now ready to insert the DNA for the secretion signal sequence, 6x His and Factor Xa into pFBΔBH1. A BssHII restriction site before the

secretion signal sequence was introduced by PCR, using pAcGP67-TfR as template. The reverse primer spans the EcoR1 site between the 6x His tag and $TfR^{121-760}$. These are the sequences, in the 5' to 3' convention, of the primers used:

Forward: AAT GCG CGC ATG CTA CTA GTA AAT CAG

Reverse: CGC GAA TTC ACC ACG TCC CTC GAT

Due to incompatibility in optimum temperature, we sequentially digested both the plasmid (pFBΔBH1) and the PCR product with BssHII followed by EcoR1. For the plasmid it is important that the order indicated be followed because BssHII is not an efficient cutter when the site is close to the end of the DNA, while EcoR1 remains efficient.

After purifying the digested materials by agarose gel electrophoresis, ligation reactions were set up following the protocol below. The molar ratios are pFBΔBH1:insert. A 1:1 molar ratio contains approximately 26μg pFBΔBH1/μg insert. The concentrations of the starting materials are 250 ng/μl for pFBΔBH1and 4.5ng/μl for insert, as determined from O.D at 260 nm.

Molar Ratio	dd-pFBΔBH1	dd-insert	T4 buffer	T4 DNA ligase	Water
1:1	4	0.9	2	0.5	12.6
1:3	4	2.6	2	0.5	10.9
1:10	4	8.6	2	0.5	4.9
Vector only(1)	4	0	2	0.5	13.5
Vector only(2)	4	0	2	0	14
Insert only	0	2.6	2	0.5	14.9

Table 1. Ligation protocol to insert the gp67 secretion signal sequence and 6x His tag into pFBΔBH1. Volumes of each component are in microliter. dd-doubly digested

The ligation was allowed to proceed overnight at 4°C. MC1061 competent cells were transformed with the ligated products. A total of five clones were selected from different ampicillin plates, grown for twelve hours at 37°C in 2 ml Luria broth with ampicillin, and plasmid DNA was purified from each clone. To determine if the desired fragment was successfully inserted, 15 μl aliquot from each prep was digested with BssHII and EcoR1. An equivalent amount of pFBΔBH1 was treated the same way as a negative control. Figure 9 clearly shows that the insert is present in all clones isolated.

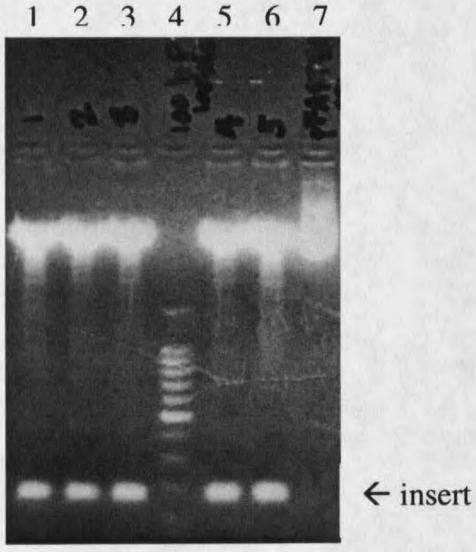

Figure 9: Lanes 1,2,3,5 and 6 are all clones that were isolated. Lane 4 contains the 100 bp ladder and lane 7 contains a negative control

With the secretion signal sequence, 6x His Tag and Factor Xa cleavage site in pFBΔBH1, now referred to as pFBΔBH1-gp67, we are now ready to incorporate the $TfR^{121-760}$ cDNA into the transfer vector. Although it was possible to have just PCR amplified the entire plasmid to do a single step incorporation of the insert into the new vector, we decided to take the longer path to avoid the risk of having PCR errors and the

need to sequence a 2 kb insert to validate our experiment. To excise the TfR-coding fragment from pAcGP67-TfR, 10 μg of the plasmid was digested with EcoR1 and Not1. pFBΔBH1-gp67 was also doubly digested with EcoR1 and Not1 to generate sticky ends complementary to those of the TfR-coding fragment from pAcGP67-TfR. The linearized vector, pFBΔBH1-gp67, and the TfR cDNA were both extracted by standard procedures from agarose gel-separated bands. Stock solutions were made with concentrations of 175 ng/μl for the TfR cDNA and 50 ng/μl for doubly digested pFBΔBH1-gp67. The ligation protocol was as follows:

Molar Ratio	dd-pFBΔBH1-gp67	TfR cDNA	T4 buffer	T4 DNA ligase	Water
1:1	4	0.5	2	0.5	13
1:3	4	1.5	2	0.5	12
1:10	4	5	2	0.5	8.5
Vector only(1)	4	0	2	0.5	13.5
Vector only(2)	4	0	2	0	14
Insert only	0	1.5	2	0.5	16

Table 2. Ligation protocol to insert TfR cDNA into pFBΔBH1-gp67.

These reactions were allowed to proceed overnight at 4°C. 100μl aliquots of freshly thawed MC1061 competent cells were transformed with 5 μl from one ligation condition and allowed to grow and express the selectable marker in 2 ml Luria broth. 100-, 1,000- and 10,000-fold dilutions were made for each and 150, 100 and 50 μl from the respective dilutions was streaked on ampicillin-containing agar plates. Colonies that grew in each of the 10,000-fold dilute samples were counted.

Description	Number of Colonies
1:1	4
1:3	4
1:10	2
Vector only(1)	0
Vector only(2)	1
Insert only	0

Table 3. Comparison of the number of transformants grown in Amp-containing agar plates.

Ten colonies were picked for further screening. Plasmid DNA was prepared from each colony and a triple digest with BssHII, EcoR1 and Not1 was carried out to determine the presence of the TfR-coding fragment and the fragment that codes for the gp67 secretion signal sequence, 6x His and Factor Xa cleavage site. The picture below shows the presence of all the expected fragments in seven out of ten clones that were picked. We refer to this construct as pFBΔBH1-gp67-TfR$^{121\text{-}760}$.

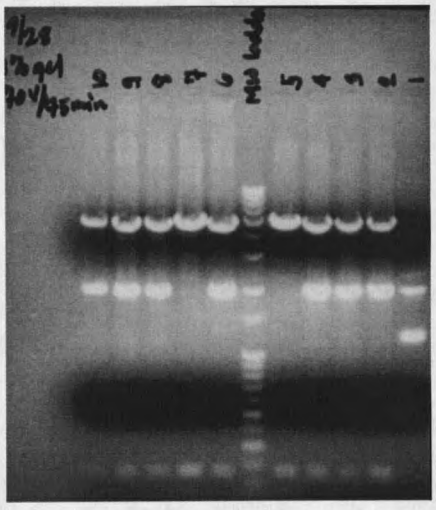

Figure 10. Triply digested plasmid DNA from isolated clones

Although we know that the complete insert we've made is identical to that of the source plasmid, pAcGP67-TfR, the protein they've isolated from this construct does not crystallize, at least not without another macromolecule. Removing the 6x His tag by Factor Xa digestion leaves three more residues (GFF) attached to the N-terminal of the ectodomain of TfR. We suspect that, aside from the difference in the host organism for overexpression, these extra residues drastically alter the crystallization conditions, because we know that the N-terminus is involved in crystal contacts.

The ectodomain of TfR, comprised of R121 to F760, is released when TfR-containing membranes are treated with trypsin. Trypsin cleaves between R120 and R121 but it is possible that the N-terminal residues may be important for recognition of this site. Thereafter, we decided to incorporate residue P117 to R120 into our construct, allowing Trypsin cleavage to produce TfR 121-760.

To do this we searched for an enzyme that cuts in the TfR-coding region but not anywhere else in the entire vector. We identified Nde1 as a single cutter that, together with EcoR1, will cut out a 98 base pair fragment from pFBΔBH1-gp67-TfR$^{121\text{-}760}$ that spans R121. This was replaced with a PCR amplified fragment from pCMVTfR, a plasmid used to overexpress TfR$^{117\text{-}760}$ in CHO cells that contained residues P117-R120. The primers used to amplify the fragment of interest from pCMVTfR are:

Forward primer: ATA GAA TTC CCT GCA GCA CGT CGC

Reverse primer: CGC ATC TTT TTG AGA TCC AGC CTC

We introduced an EcoR1 cleavage site before P117 in the forward primer. The reverse primer includes 33 bases more after the Nde1 site because Nde1 does not cut as well near the ends.

pFBΔBH1-gp67-TfR$^{121-760}$ and the PCR amplified fragment from pCMVTfR were doubly digested with EcoR1 and Nde1 and purified by standard techniques. A similar ligation protocol as those enumerated in table 1 was followed and five colonies were picked from the transformants. Plasmid DNA was purified from minipreps and digestion of small aliquots from each clone with EcoR1 and Nde1 confirmed the presence of the low molecular weight insert, as shown in Figure 11. This version was called pFBΔBH1-gp67-TfR$^{117-760}$. This plasmid, together with pFBΔBH1-gp67-TfR$^{121-760}$, was submitted for sequencing, and results confirmed the presence of the right insert.

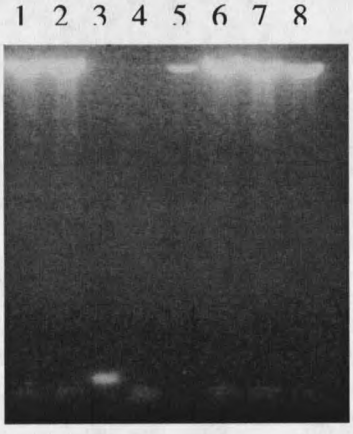

Figure 11: Lanes 1, 2, 6, 7, 8 are clones that the contain the right insert. Lane 3 is an aliquot of the undigested PCR product, that's why it is heavier. Lane 4, is some doubly digested PCR product and lane 5 is doubly digested pFBΔBH1-gp67-TfR$^{121-760}$.

CHAPTER 3

Generation of Recombinant Baculovirus and Overexpression of TfR from Sf9 Cells

Following the manufacturer's protocol, we transformed DH10Bac cells with pFBΔBH1-gp67-TfR$^{121-760}$ and pFBΔBH1-gp67-TfR$^{117-760}$. Cells that contain the recombinant bacmid are white in the presence of a chromogenic substrate, Bluo-gal, in contrast to blue colonies containing the wild type bacmid. White colonies were selected and restreaked for validation. One can clearly see in Figure 12 the blue colonies that contain the unaltered bacmid. The fewer white colonies the transposition of the gene of interest into the plasmid disrupted expression of the lacZα peptide. White colonies contain recombinant bacmids.

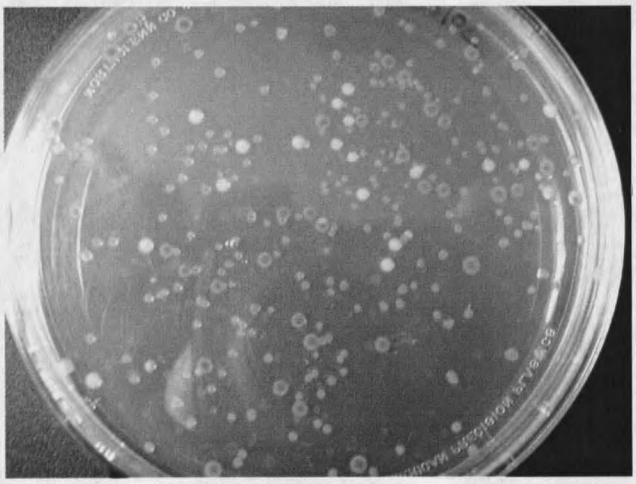

Figure 12 : Agar-plate with transformed DH10Bac *E.coli* cells (cell dilution 10^{-3}) showing blue and white colonies.

Using reagents from the Qiagen miniprep kit, high molecular weight bacmid DNA of each recombinant clone was isolated and the success of the site-specific transposition was confirmed by PCR. An agarose gel of the amplified fragments is shown in Figure 13. The smaller fragment is approximately 2 kbp. It is a portion of the recombinant bacmid that is amplified by a forward primer (5'-CGCCAGGGTTTTCCCAGTCACGAC-3') that anneals at a site upstream of the multiple cloning site of pFastBac1 and a reverse primer (5'-CGCATCTTTTTGAGATCCAGCCTC-3') that partly anneals at the region of TfR that codes residue E156 to D162. The larger fragment is approximately 3kbp. It contains the entire insert (gp67 secretion signal sequence-6xHis tag-Factor Xa site-TfR117-760), which is about 1900 bp, and another 1kb portion that is a part of pFastBac1 plasmid. The primers used to amplify the heavier fragment are:

Forward: 5'-AAT GCG CGC ATG CTA CTA GTA AAT CAG-3'

Reverse: 5'-AGC GGA TAA CAA TTT CAC ACA GGA-3'

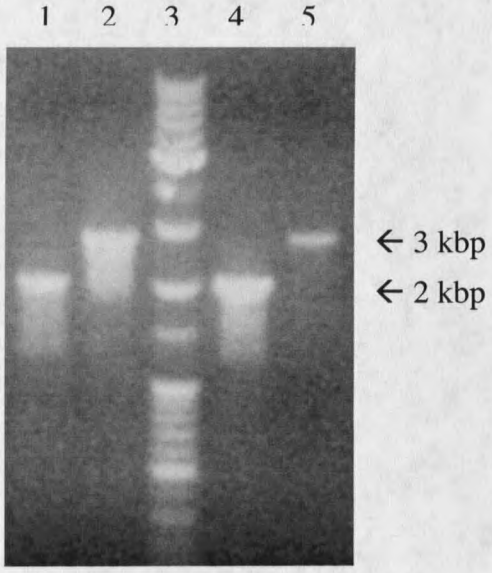

Figure 13: Agarose gel of PCR amplified portions of the recombinant bacmid. Lanes 1 and 2 -- pFBΔBH1-gp67-TfR$^{117-760}$; lanes 4 and 5 -- pFBΔBH1-gp67-TfR$^{121-760}$; lane 3 -- mixture of 100 bp and 1 kbp MW ladders.

To generate primary viral stocks of each construct, 5μL aliquots of the bacmid DNA was used to transfect adherent Sf9 cells, following manufacturer's protocol. After 72 hours of incubation at 27±1°C in an incubator, the supernatant was collected and clarified by centrifugation at 500 x g for five minutes. 45 μl of each clarified supernatant was loaded on an 8% SDS-PAGE gel. The bands were transferred to nitrocellulose for western blot analysis using mouse anti-6x His primary antibody and rabbit anti-mouse Ig antibody conjugated with alkaline phosphatase as secondary antibody. The blot for TfR$^{117-760}$ is shown as Figure 14 below, while that for TfR$^{121-760}$ did not show any sign of expression. At this point we decided to use TfR$^{117-760}$ for all succeeding work.

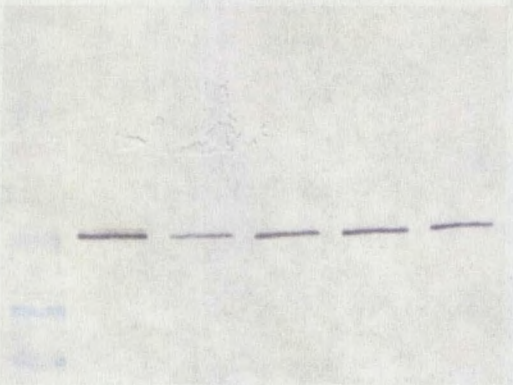

Figure 14. Western blot confirming the presence of His-tagged TfR.

Assuming an initial titer of $2x10^7$pfu/ml of the primary viral stock obtained above, a 50 ml suspension culture of Sf9 cells at mid log phase was infected at a multiplicity of infection (MOI) of 0.1. The amount of virus required was calculated as follows:

$$\text{Inoculum required (ml)} = \frac{\text{MOI (pfu/cell)x(total number of cells)}}{\text{Viral titer (pfu/ml)}}$$

The culture was maintained in the incubator at 27±1°C with stirring rate between 80 to 100 rpm. This secondary viral stock was harvested after 48 hours by centrifugation at 500x g for five minutes. A titer between 1 to $1.5x10^8$ pfu/ml was typically obtained at this step and this stock was used for large-scale viral amplification (>1L).

MOI Optimization and
Small Scale Protein Expression

To determine the optimum MOI for protein expression and the time required to obtain the best quality of protein we carried out a time course experiment at four different MOI: 1, 2, 5, 10. Four 100 ml spinner flasks containing 50 ml suspension culture with